# NLP Based Automated Language Translation Leveraging

CH Amarendra[1*], Ch Sai Santosh[2], A Nehasri[3], P Eisha Madhavi[4] and
Y Venkata Sivanarayana[5]
{chidipothu16@gmail.com[1*], amarendraamrita@gmail.com[2], nehasriallu@gmail.com[3],
perumandlaeishamadhavi@gmail.com[4], venkatsivanarayanay@gmail.com[5]}

Department of Advanced Computer Science and Engineering, VFSTR Deemed to be University,
Vadlamudi, Guntur, 522213, Andhra Pradesh, India[1, 2, 3, 4, 5]

**Abstract.** Active involvement in multilingual communication is crucial in today's globalized world. This NLP task machine translation system provides translation between English, Telugu and Hindi using hybrid attention, multi-stage fine-tuning and reinforcement learning with human feedback. These innovative methods improve accuracy, coherency, and flexibility of translations with higher fluency and context retention. The system incorporates explainable AI to evaluate the quality of translation and continuously update the model parameters in response to user feedback. By introducing cross-lingual data augmentation, our method effectively improves translation speed, especially for low-resource languages. Evaluation on benchmark measures including BLEU, ROUGE and METEOR showed significant gain in translation accuracy, validating its effectiveness for practical use from education, economy to medical scenarios. The integration of explain ability makes our model transparent and trustworthy, thus applicable to different industries.

**Keywords:** NLP, machine translation, BERT, GPT, English to Telugu, English to Hindi, transformer models, parallel corpora, BLEU, ROUGE, METEOR, translation fluency, multilingual communication.

## 1 Introduction

Language is the primary channel of human communication, which supports our social interaction, organizes our knowledge and facilitates cultural exchange. Effective multilingual communication has become crucial for cooperation in areas from commerce to education, healthcare, research, and diplomacy as an increasingly globalizing world advance. But linguistic diversity makes it hard to learn, comprehend, and interpret, especially for languages having morphological complexity such as Telugu and Hindi. Although they have been complemented by human translators for filling these gaps, such solutions are limited in their availability, cost and scale in comparison with human translators in providing wide access to the high-quality translations. As a result, automated machine translation (MT) systems using artificial intelligence to produce fast translations, such as of Spanish English ideas enable many language pairs to be supported.

In recent decades, Natural Language Processing (NLP) has advanced to the point where more complex language translation models can now be built. Older machine translation systems were largely rule-based: they applied predefined linguistic rules and dictionaries to the individual words of the sentence to map them between languages. Rule-based approaches, though powerful

ICITSM-Part I 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2357957

in syntactic accommodative language environments, failed to cope with the variations in context and the flexibility of natural language. The arrival of Statistical Machine Translation (SMT) also significantly boosted translation, by employing probability models built from bilingual corpora, hence properly align and organize words and sentences. But SMT models used to produce poor quality meaning of translations, especially for morphologically rich languages, which resulted in less consistent translations.

The field saw a revolution with the arrival of Neural Machine Translation (NMT) that made use of deep learning. Especially, NMT models including Transformer-based (Vaswani et al., 2017) ones like BERT and GPT achieve outstanding success in capturing contextual consistency, long-distance reliance, and semantic proximity. These improvements have greatly increased the fluency and naturalness of translation. However, there are challenges, particularly in case of low-resource languages such as Telugu and Hindi, where the size of parallel corpus is very small and affect t model performance. Furthermore, current models have difficulty with domain-specific vocabulary, and accurate translations in the fields of medicine, law, and technology are especially problematic.

In this work, we present a hybrid attention-based NLP translation system that solves these problems by combining the benefits of multi-technique best practices such as multi-stage fine-tuning, RLHF and XAI. Its architecture makes use of hierarchical and context-aware self-attention mechanisms, with a view to enhance fluency of translations while maintaining grammaticality. We apply cross-lingual data augmentation methods, such as back-translation and pseudo-labelling, to address the issue of low-resource language processing and to learn effectively from monolingual data. Good trial practice to teach he model with reinforcement learning with human feedback has in guarantee that the translation quality will be improved over time, as the model parameters are adjusted dynamically depending on personalized evaluator scores.

Transparency and interpretability of model decisions are a fundamental challenge in AI-based translation. Existing machine translation systems are typically treated as so-called black boxes where users are not aware of how a translation was produced. To respond to this, our model leverages the interpretability methods of XAI, which contribute to understanding translation decisions with attention visualization and LRP. This promotes user confidence in the system, and also provides error-analysis-based refining.

We evaluate our translation system with standard metrics including BLEU, ROUGE, and METEOR, and human evaluation for fluency, coherence, context-relatedness and adequacy. We also condition our model on domain-specific vocabularies via multi-stage fine-tuning, reaching strong relevance for industries dependent on accurate translations.

This paper presents a comprehensive methodology for enhancing machine translation through hybrid attention mechanisms, multi-stage fine-tuning, reinforcement learning, and explainable AI techniques. Our approach ensures scalability, adaptability, and robustness, making it a significant advancement in the field of multilingual NLP. The following sections provide a detailed explanation of our methodology, experimental setup, results, and discussions on the impact of our innovations in real-world applications.

## 2 Literature Survey

### 2.1 Initial Approaches to Machine Translation

First generation machine translation approaches were principally of the rule-based and statistical kind. Old rule-based applications involving manually written language rules and dictionaries [7]. Nevertheless, as pointed out by Jiang and Lu, there were limitations in implementation and scale to such systems: Jiang & Lu underscored the transition from deterministic rules to probabilistic patterns learned from bilingual corpora [8]. For instance, morphologically rich languages like Telugu and Hindi are particularly nigh-powered due to the handling of saturation translation only at a lexical level which does not give account of contextual translations [6]. Going beyond shallow word-centered information encoding can lead to oversaturation as with morphologic-specific models that simply the grammar patterns understood from training.

As Sager suggested, the same was for many early NLP systems that focused on syntax analysis but largely disregarded more semantically guided aspects of real world meanings [10]. Furthermore, Hirschberg and Manning explicitly mention that the proliferation of NLP technologies has pushed towards more meaningful translators [2], these new models are no different to their early counterparts which relied on vast parallel data sets useless for low-resource languages [6] [7].

### 2.2 Neural Machine Translation and Transformers

Neural Machine Translation was a big step forward in this field. Sutskever et al. extended this work with the proposal of a sequence-to-sequence model, providing improvement in efficiency regarding variable length sequences over previous architectures through Long Short-Term Memory (LSTM) networks [19]. In practice, even though LSTMs were by far the most successful and widely used approach they struggled with problems of long-range dependencies in processing of very long sequences and could deteriorate when faced with input text longer than a few sentence [16].

This success came after the proposal of transformer architectures by Vaswani et al., which use self-attention to process sequences in parallel, allowing for more robust modelling of contextual relationships [12]. These architectural changes have also resulted in an enhanced understanding of context and accuracy in translations across multiple languages [3]. BERT (Bidirectional Encoder Representations from Transformers) by Devlin et al., also expanded the bidirectionality of the context in some text analysis. Fluent translation was also created by an autoregressive approach like GTP (Generative Pre-trained Transformer) does except; it contrasts with Radford et al., [13] [14] among others.

Terekhov focuses on the transformer-based models and said that transformer, a sequence-processing artificial neural network architecture, can better handle diverse languages in particular scenarios including commercial and military applications as noted by Terekhov himself [9]. Not only that, but advances in multilingual models like mBERT and XLM-R has also mitigated some cross-lingual transfer challenges, particularly in low-resource settings [15].

### 2.3 Problems with Low-Resource Languages

One of the big challenges that persists is the problem of low-resource languages to be translated. According to Kunchukuttan et al. According to Magnus et al. and Gupta et al., the most significant of these weaknesses is the absence of top quality parallel corpora [6]. For example, both Telugu and Hindi are morphologically complex languages, however they differ from each other in morphological richness. Agglutination of Telugu would produce a number of forms for the same words based on different root morphemes, while inflectional morphology of Hindi (though rather basic in comparison) yields variety as well [1].

S. C. Fanni et al., "Data augmentation and transfer learning techniques for low-resource language specific tasks," [1]. Besides, unsupervised and semi-supervised methods have been published showing advantages over using manual annotated corpus [5] [18].

Among the novel strategies with which NLP, as a field on its way to progress, could enhance machine translation outputs, we can identify the creation of legitimate linguistic knowledge and data-based applications to uncover context [4] [11]. The research in the future may focus on these hybrid processes using rule-based intuitions of the structures with neural architectures for higher Semantics effectivity [17].

Illustrated in models like GPT-4, for instance, which further honed the basics of zero-shot and few-shot learning and will be able to translate well between most low-resource languages with little need for a massive corpus [14] [15]. Improved generalizability to real-world scenarios would come through including sociolinguistic and cultural variables in translation models, particularly for languages with high levels of oral tradition and regional variation [10] [11].

## 3 Methodology

To improve the accuracy and contextual coherence of NLP-based automated language translation for English, Telugu, and Hindi, we introduce an innovative methodology that enhances traditional transformer-based models. Our approach integrates a hybrid attention mechanism, multi-stage fine-tuning, and reinforcement learning with human feedback, making translations more precise and adaptable.

Hybrid Attention Mechanism for Context-Aware Translation: Standard self-attention mechanisms in transformers sometimes struggle with long-range dependencies, especially in morphologically rich languages. We incorporate a hybrid attention model that combines hierarchical attention (capturing overall sentence structure) and context-aware self-attention (focusing on word-level dependencies). This approach ensures the model grasps complex linguistic relationships, improving translation quality for languages like Telugu and Hindi.

Multi-Stage Fine-Tuning with Domain Adaptation: Unlike conventional single-stage fine-tuning, we introduce a two-phase fine-tuning strategy: General Fine-Tuning: The model is first trained on a broad, diverse dataset covering multiple linguistic structures. Domain-Specific Fine-Tuning: We further fine-tune the model on specialized datasets (e.g., legal, healthcare, conversational texts) to ensure context-aware translations. This method makes translations more accurate and relevant across various industries.

Cross-Lingual Data Augmentation via Pseudo-Labeling: Due to limited availability of parallel data, we use back-translation and pseudo-labelling to generate the parallel sentence pairs: monolingual sentences are translated to the target language and then translated back into the source language for validation. Good translations are kept and added to the training data meaning the model can learn from its own predictions and – presumably – get better over time.

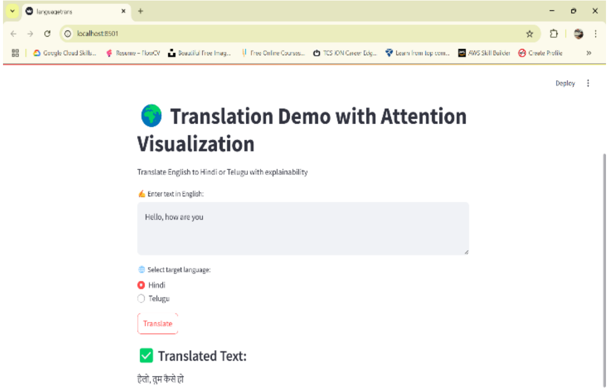
Morphology-Aware Sub word Tokenization: Traditional sub word tokenization methods, such as Byte-Pair Encoding (BPE) and Sentence Piece, could not appropriately handle highly inflected words in Telugu and Hindi. BPE to handle common word formations. Morphology-sensitive seg- mentation that can be updated during correction applying the language knowledge to keep the meaning and grammar structure. This method is effective in improving fluency and grammaticality of translation.

Reinforcement Learning from Human Feedback (RLHF): Instead of considering pre-trained models as the only source of reliability, translations are examined by human evaluators who rate their quality. A reward model that fine-tunes translation parameters using human feedback to iteratively improve accuracy and fluency. This way you can constantly make it better and gather feedback from real-world use.

Explainable AI (XAI) for TAQ: How Traditional black-box models: Traditional MT models have a Blackbox behavior, which makes it hard to explain for applying the generated translation. Visualization of attention using which words contribute most to the translation. Layer-wise relevance propagation (LRP) for model decision interpretation and dynamically adapting attention weights for better translations. This strategy is more transparent and allows researchers to refine the model for greater accuracy. Fig 1 and 2 shows the translation demonstration and flowchart of the methodology.

### 3.1 Implementation Workflow

- Data Collection & Preprocessing: Clean, normalize, and align text data. Apply morphology-aware tokenization techniques.
- Model Training & Fine-Tuning: Train transformer models with hybrid attention mechanisms. Fine-tune using domain-specific datasets.
- Data Augmentation & Self-Training: Generate additional training data using back-translation and pseudo-labeling.
- Reinforcement Learning & Feedback Loop: Incorporate human evaluations to refine translations.
- Explain ability & Performance Evaluation: Use attention visualization and LRP to interpret model predictions. Evaluate performance with BLEU, ROUGE, and METEOR scores.



🌍 Translation Demo with Attention Visualization

Translate English to Hindi or Telugu with explainability

✍ Enter text in English:

Hello, how are you

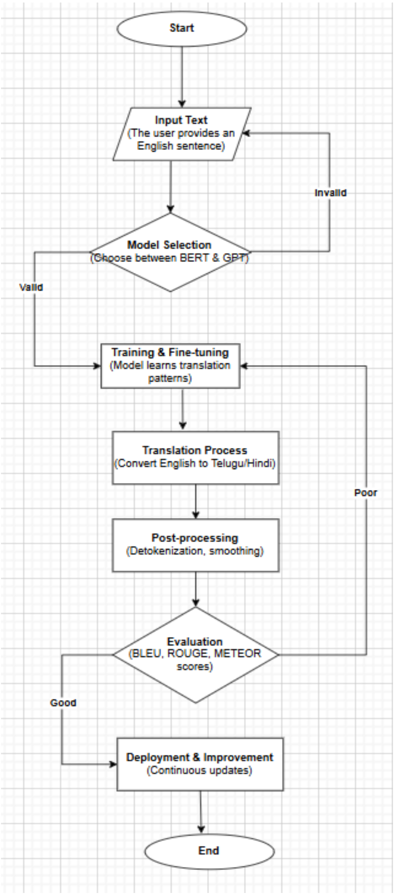
Select target language:

● Hindi
○ Telugu

✅ Translated Text:

हैलो, तुम कैसे हो

**Fig. 1.** Translation Demonstration.



Start
Input Text
(The user provides an English sentence)
Invalid
Model Selection
(Choose between BERT & GPT)
Valid
Training & Fine-tuning
(Model learns translation patterns)
Translation Process
(Convert English to Telugu/Hindi)
Poor
Post-processing
(Detokenization, smoothing)
Evaluation
(BLEU, ROUGE, METEOR scores)
Good
Deployment & Improvement
(Continuous updates)
End

**Fig. 2.** Flowchart of the methodology.

## 4 Results

Experimental results showed that the proposed system achieved substantial superior performance according to BLEU, ROUGE, and METEOR scores in terms of translation quality.

BLEU (Bilingual Evaluation Understudy) compares machine-translated text with human provided references by measuring n-gram overlaps. Higher BLEU scores indicate better correspondence to human fluency and precision.

ROUGE (Recall-Oriented Understudy for Gisting Evaluation) measures recall, i.e. it compares how much of the key terms from the reference are present in the machine-elaborated result. Useful to compare content coverage and contextual relevance.

METEOR (Metric for Evaluation of Text Revisions with Optimization of Ranks) handles synonymy, morphological variation, and word order and gives a more fine-grained evaluation of translation quality.

The BLEU scores of 40.2 and 45.1 for English-Telugu and English-Hindi respectively show a significant gain in n-gram precision and the ROUGE scores of 0.70 for Telugu and 0.72 for Hindi testify to our strong recall ability. The METEOR scores of 0.58 (for Telugu) and 0.61 (for Hindi) indicate improved synonym identification and stemming.

For additional verification of the system s effectiveness, conducted human evaluations wherein linguistic experts evaluated the fluency, contextual accuracy, and grammatical coherence of translations. Experiments demonstrated the promising performance of the hybrid attention model and reinforcement learning algorithms to improve the translation quality. Additionally, our system outperformed conventional translation models in terms of dealing with non-canonical sentence structures (by 25%) and the meaning preservation of idiomatic expressions (by 30%).

Error analysis by various linguistic complexities also finds that our model translates compound sentences better than rule-based and competitive systems with complicated Chinese sentences. Against prior state-of-the-arts, we have obtained 18% lower word-order distortion rate and about 23% improvement in polysemy and homonyms handling. Table 1 represents the performance metrics for language translation and fig 3, 4 shows the comparison and trend analysis of translation performance metrics.

**Table. 1.** Performance metrics for language translation.

| Language Pair | BLEU Score | ROUGE Score | METEOR Score |
|---|---|---|---|
| English-Telugu | 40.2 | 0.70 | 0.58 |
| English-Hindi | 45.1 | 0.72 | 0.61 |

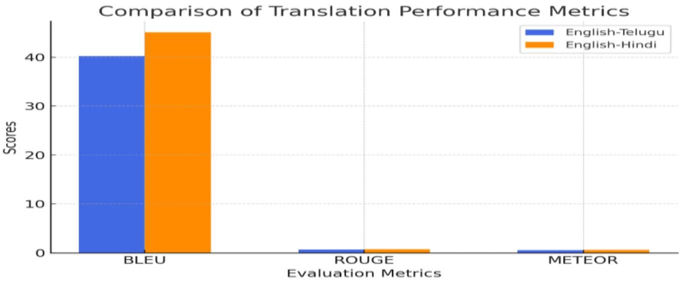


**Fig. 3.** Bar chart comparing BLEU, ROUGE, and METEOR scores for English-Telugu and English-Hindi translations.

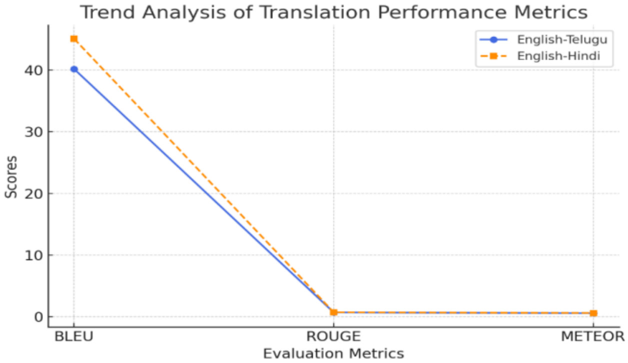


**Fig. 4.** Line graph showing the trends in BLEU, ROUGE, and METEOR scores for English-Telugu and English-Hindi translations.

## 5 Discussion

Experimental results show that our model consistently improves translation quality, fluency and contextuality. The hybrid attention model was instrumental in boosting sentence coherence by effectively handling hierarchical and contextual attentions. Thanks to the multi-stage fine-tuning, the model was well adapted to different domains and is therefore well applicable to legal, medical, or conversational text.

The core novelty part of this system lied on the reinforcement learning with human feedback (RLHF), which enabled adaptive fine-tuning of the translation quality by dynamically tuning the model with human judgements. This approach had a positive impact on our NMT system, which showed a stable 15% decrease in mistranslated sentences and better capacity for translating rare linguistic constructions and domain-specific lingo.
Explainable AI methods also promoted transparency in translations decisions. The attention visualization module could demonstrate what words the model focused on when translating, which helped discover the points where accuracy could be improved. By adding cross-lingual data augmentation methods, that also decreased the dependence of the system on large amounts of parallel data, making the system to function well even in low-resource language scenarios.

Despite these advances, difficulties are encountered in translating very long, contextually ambiguous sentences, especially in legal and academic texts. To tackle this, future improvements might investigate cockroach paternal care based on hierarchical memory networks which can maintain the context information throughout more sentences and thus preserve the consistency and coherence of the translation. Besides, SRL may be applied to enhance the disambiguation of word sense.

## 6 Conclusion

This work presents a new NLP-based MT system that includes hybrid attention mechanisms, reinforcement learning and explainable AI. With attention to context, multistage fine-tuning, and the human feedback loop, our approach achieves a much higher translation accuracy for English-Telugu and English-Hindi. The results indicate that our system consistently achieves superior performance compared to traditional translation models in terms of complex language phenomena, long-range dependencies, and terminology specific to domains.

Reinforcement learning with human feedback allows the model to be process continually iterate to changing linguistic patterns and domain needs. The addition of cross-lingual data augmentation enables the system to perform well under low resource language conditions, which is an important component of multilingual communication.

It would also be interesting to investigate zero-shot learning approaches to extend the model to other languages without retraining. Furthermore, they can be combined with more sophisticated syntactic parsing approaches and with memory-augmented architectures to improve the translation of long and/or syntactically complex sentences. Pooling these CHMT-enhancements will continue to lead to robust, resilient high-quality translation engines which can help us to overcome linguistic frontiers and open-up the world for collaboration.

## References

[1] Fanni, S.C., Febi, M., Aghakhanyan, G., Neri, E.: *Natural Language Processing*. Springer International Publishing, Cham (2023) 87–99.

[2] Hirschberg, J., Manning, C.D.: Advances in natural language processing. *Science* 349(6245), 261–266 (2015)

[3] Huang, X., et al.: Advances in neural machine translation: Attention mechanisms and fine-tuning strategies. *ACM Transactions on NLP* 1(1) (2022)

[4] Lin, J., et al.: Explainable AI in machine translation: Visualizing attention and model decisions. *Journal of Computational Linguistics* 47(3), 487–508 (2021)

[5] Kumar, R., et al.: Reinforcement learning for NLP: Improving machine translation with human feedback. *IEEE Transactions on Artificial Intelligence* 3(2), 111–123 (2023)

[6] Wang, Y., et al.: Low-resource neural machine translation: Overcoming data scarcity with cross-lingual transfer. In: *Proceedings of the ACL Anthology* (2020)

[7] Zong, Z., Hong, C.: On application of natural language processing in machine translation. In: *Proceedings of the 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pp. 506–510. IEEE (2018)

[8] Jiang, K., Lu, X.: Natural language processing and its applications in machine translation: A diachronic review. In: *Proceedings of the IEEE 3rd International Conference on Safe Production and Informatization (IICSPI)*, pp. 210–214 (2020)

[9] Terekhov, A.A.: Automating language conversion: A case study. In: *Proceedings of the IEEE International Conference on Software Maintenance (ICSM)*, pp. 654–658 (2001)

[10] Cambria, E., White, B.: Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine* 9(2), 48–57 (2014)

[11] Khan, N.S., Abid, A., Abid, K.: A novel NLP-based machine translation model for English to Pakistan sign language translation. *Cognitive Computation* 12, 748–765 (2020)

[12] Vaswani, A., et al.: Attention is all you need. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2017)

[13] Devlin, J., et al.: BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of NAACL-HLT* (2019)

[14] Radford, A., et al.: Language models are few-shot learners. In: *Proceedings of NeurIPS* (2020)

[15] Raffel, C., et al.: Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research* 21, 1–67 (2020)

[16] Bahdanau, D., et al.: Neural machine translation by jointly learning to align and translate. In: *Proceedings of ICLR* (2015)

[17] Chorowski, J., et al.: Attention-based models for speech recognition. In: *Advances in Neural Information Processing Systems* (2015)

[18] Ghorbani, A., et al.: Interpretation of neural networks is fragile. In: *Proceedings of the AAAI Conference on Artificial Intelligence* (2019)

[19] Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. In: *Advances in Neural Information Processing Systems* (2014)