Journal of Computer Science

Research Article

Hindi Poetry Translation Using Neural Machine Translation

Pragya Tewari and Anurag Singh Baghel

School of ICT, Gautam Buddha University, Greater Noida, India

Article history
Received: 16-05-2025
Revised: 19-07-2025
Accepted: 07-08-2025

Corresponding Author:
Pragya Tewari
School of ICT, Gautam Buddha University, Greater Noida, India
Email: pragya.dwivedi@gmail.com

Abstract: Neural Machine Translation (NMT) has become an essential tool in Natural Language Processing (NLP), enabling the automatic translation of text across languages. However, translating poetry remains a relatively unexplored and complex task. Unlike general text, poetry carries multiple layers of meaning, emotion, rhythm, and cultural nuance, making it difficult to capture in another language through standard translation techniques. This study focuses on the translation of Hindi poetry into English using modern NMT approaches. The goal is to make the works of celebrated Indian poets more accessible to non-Hindi-speaking audiences while preserving their poetic qualities. We review existing poetry translation systems, discuss the challenges they face, and explore how translation quality is typically measured. We implemented two NMT models, an attention-based recurrent neural network and a transformer architecture, and fine-tuned them using a curated Hindi poetry dataset. Evaluation results show that our models significantly improve translation quality compared to commonly used online tools, with up to a 15% increase in BLEU scores. In addition, we consider how newer AI techniques, such as Retrieval-Augmented Generation (RAG), could further enhance poetry translation by providing contextual and cultural information during generation. Our work highlights the need for translation systems that go beyond literal meaning and better capture the expressive nature of poetic language.

Keywords: Neural Machine Translation, Hindi Poetry, LSTM Encoder-Decoder, Transformer Models, Evaluation Metrics

Introduction

Recent developments in natural language processing (NLP), especially those driven by deep learning, have brought major improvements across multiple language tasks. Within machine translation, Neural Machine Translation (NMT) has emerged as a leading method, attracting significant interest from both research and industry communities. While the capabilities of NMT have grown steadily, its application to artistic domains like poetry remains less explored.

Poetry translation presents unique challenges, as it requires not only transferring meaning but also preserving form, emotion, and stylistic elements. Despite recent progress in machine translation, capturing the subtleties of poetic language remains difficult. Nonetheless, the global circulation of poetry continues, highlighting the need for improved approaches. This study investigates how neural models can be applied to this task.

Machine Translation (MT) refers to the use of automated systems to convert text from one language (e.g., Hindi) into another (e.g., English), while retaining both meaning and fluency. Specific challenges in Hindi--English translation include:

1. **Synonym Variability:** Hindi often offers multiple synonyms for the same idea, which complicates training. Our method aims to treat synonymous words uniformly.
2. **Word Form Variations:** The same word may appear in visually different forms, such as सुन्दर and सुंदर.
3. **Semantic Shifts:** Even small synonym replacements can change sentence meaning due to Hindi's grammatical dependencies. For instance, अवकाश होता है versus छुट्टी होती है.

Many existing translation tools struggle to produce meaningful Hind-English poetic translations. They often miss contextual cues and stylistic nuances. In this paper, we address these challenges and analyze current translation systems focused on poetry. Our goal is to improve translation quality using custom trained NMT models. Our experiments suggest that these models perform better than off the shelf online translators when adapted for poetry.

DOI: 10.3844/jcssp.2026.360.366

Background

At its core, a poem or verse (कविता/पद्य) is a literary form that uses language to convey emotion, whether gentle or intense, depending on the context. Poetic language is often chosen for its aesthetic sound and expressive power, rather than for strict grammatical rules. This sets poetry (पद्य) apart from prose (गद्य). A wellknown Hindi children's poem illustrates this:

"मछली जल की रानी है,
जीवन उसका पानी है |
हाथ लगाओगे डर जाएगी,
बाहर निकालो मर जाएगी |"

Features of Poetry

Understanding the core features of poetry is essential when attempting to translate it, particularly across languages. Key characteristics include:

Verse (काव्य) and Free Verse (Mukta-Kavita मुक्त-कविता)

Poems are generally categorized into structured verse (काव्य), which follows consistent patterns, and free verse (मुक्त-कविता), which allows greater flexibility.

Syllable (Maatraa मात्रा)

Each character (अक्षर) in Hindi contributes a specific syllable value (मात्रा). Common values include 0, 1, or 2, depending on the combination of letters and diacritics.

Rhyme (Tuk तुक)

Rhyme involves repeating similar sounding words, especially at the end of lines. In Hindi, rhyming often requires consistent syllable counts across lines.

Rhythm (Laya लय)

Rhythm refers to the repeated patterns in a poem's phrasing, creating a musical quality and natural flow. It influences where pauses and emphases occur.

Meter

Meter relates closely to rhythm and syllables, defining the length and beat structure of each poetic line.

Form

Poems may vary in rhythm or rhyme, yet still belong to distinct forms such as lyric, narrative, or descriptive.

Structure of Poems

While poetic structures vary by language, certain elements remain consistent. Hindi poems are typically composed of the following elements:

- Sthaayee (स्थायी) or tek (टेक)
- Antara (अंतरा)
- Samaapti (समाप्ति)

Supporting lines (sahyogee) may also be included. These units mirror the concept of stanzas in English poetry. Often, the antara and sthaayee differ in rhyme, but the samaapti usually follows the same pattern. All parts should maintain consistent laya (rhythm) and maatraa (syllable count).

Overview of Machine Translation Techniques

Significant advances in Machine Translation (MT) have improved cross language communication. MT systems are typically categorized as:

1. Statistical Machine Translation (SMT): SMT uses statistical models trained on bilingual corpora to map words and phrases across languages. Phrase Based SMT (PBSMT) enhances accuracy by translating in phrases rather than single words, though it can struggle with long dependencies or grammar mismatches.
2. Rule Based Machine Translation (RBMT): RBMT applies linguistic rules and dictionaries to translate sentences. While precise, it may ignore context and perform poorly on idiomatic expressions.
3. Hybrid Machine Translation (HMT): HMT combines SMT and RBMT, using translation memory and human input to improve accuracy.
4. Neural Machine Translation (NMT): NMT uses deep neural networks to model entire sentences, offering better fluency and contextual handling than previous methods.
5. Domain Adaptation: This technique adapts general MT models to specific fields (e.g., poetry), especially when training data is limited, making it valuable for low resource languages.

Related Work

Over the years, Machine Translation (MT) approaches have undergone substantial evolution, with statistical and neural models emerging as dominant paradigms. (Seljan et al., 2020) explored MT for translating poetry in a lowresource language setting (Croatian--German). Using a dataset composed of original poems and expert human translations, their study demonstrated that MT could be effective even in literary domains.

In the area of poetry generation, Wei et al. (2018) presented a two stage model that balances content planning and stylistic control. Their system used a modified recurrent encoder--decoder framework to sequentially produce lines of poetry. Similarly, Ghazvininejad et al. (2018) proposed a neural poetry translation system that maintains rhyme and rhythm by transforming source lines into structured English verse.

Wang et al. (2022) offered a broad overview of improvements in NMT, including developments in real time translation that balance quality and latency. For Mandarin poetry, Wang et al. (2016) introduced a system

DOI: 10.3844/jcssp.2026.360.366

that first plans subtopics and then generates lines using a sequential neural model.

Focusing on Hindi-English translation, Gangar et al. (2021) implemented a transformer based model trained specifically for this language pair, showing its suitability for low resource tasks.

The rise of deep learning has positioned NMT as a preferred method in MT research and applications, due to its ability to model long term dependencies and simplify architectures compared to traditional statistical methods such as PBSMT. Early NMT systems like the one introduced by Kalchbrenner & Blunsom (2013) laid the foundation for subsequent refinements (Junczys-Dowmunt et al., 2016).

More recently, generative models augmented with retrieval capabilities have emerged as a promising avenue for poetry translation. These models offer the potential to integrate contextual and cultural references, though their use in literary tasks remains at an early stage.

Materials and Methods

Proposed Approach for Hindi-English Poetry Translation Using NMT

An overview of our approach is shown in Figure 1. It depicts a deep neural network (DNN) model translating an input sentence from Hindi (source language) to English (target language).

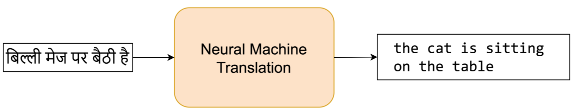

Fig. 1: Hindi to English Translation using NMT

In this section, we describe two NMT based methods used for translating Hindi poetry into English.

Encoder-Decoder Approach

We adopted the RNN based encoder-decoder model as described by Cho et al. (2014) for translation. This architecture includes two key parts: an encoder and a decoder as shown in Figure 2. The encoder processes the input sequence into a fixed size vector, which the decoder then uses to generate the translated sequence. We implemented Long Short Term Memory (LSTM) units as introduced by Hochreiter & Schmidhuber (1997) for both modules, given their strength in modeling long range dependencies in sequential data such as text and speech, as discussed by Tewari et al. (2022).

LSTM units include gates that regulate the flow of information over time. The core equations that define LSTM computation are shown below:

$$
\begin{aligned}
i &= \sigma(W^i \cdot x_t + R^i \cdot h_{t-1} + b^i)\\
f &= \sigma(W^f \cdot x_t + R^f \cdot h_{t-1} + b^f)\\
g &= \tanh(W^g \cdot x_t + R^g \cdot h_{t-1} + b^g)\\
o &= \sigma(W^o \cdot x_t + R^o \cdot h_{t-1} + b^o)\\
c_t &= f \odot c_{t-1} + i \odot g\\
h_t &= o \odot \tanh(c_t)
\end{aligned} \tag{1}
$$

Here, x_t is the input at time t, h_t the hidden state, and c_t the cell state. The symbols i, f, g, o represent input, forget, candidate, and output gates respectively. $\odot$ denotes element-wise multiplication. Weight matrices W^j, R^j, and biases b^j are all learned during training.

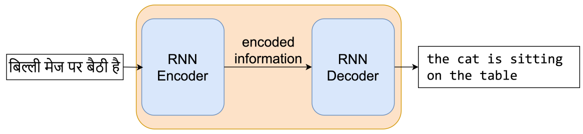

Fig. 2: LSTM encoder-decoder for Neural Machine Translation.

Each word from the input sentence is passed into the model over time steps. At each step t, the encoder updates the hidden state h_t based on current input and previous state. The encoder's final hidden vector, h_6^E, is then handed off to the decoder as its initial state as shown in Figure 3.

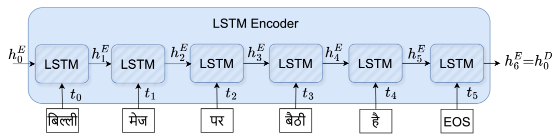

Fig. 3: Encoding the input sentence into hidden state vector (h) of the LSTM

To initiate decoding, a special <SOS> token is fed at the first time step, as shown in Figure 4. The decoder generates a probability distribution over the output vocabulary at each step, selecting the most likely word to output next.

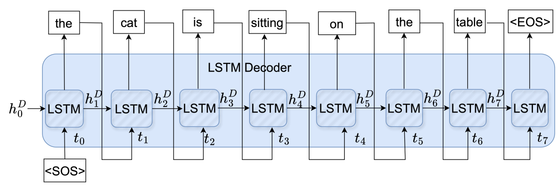

Fig. 4: Output of the decoder at different time steps

This process continues until the decoder predicts the <EOS> tag. To evaluate the accuracy of generated output, we use cross-entropy loss:

$$Loss = \sum_{w=1}^{|S|} \sum_{e=1}^{|V|} y_{w,e} \log(\hat{y}_{w,e}) \tag{2}$$

where, $|S|$ = vocabulary size, $|V|$ = sentence length, $y_{w,e} = 1$ if word e is correct at position w else 0, and $\hat{y}_{w,e}$ = predicted probability.

We further introduced a global attention mechanism (Luong et al., 2015) to enhance translation quality. This allows the decoder to selectively focus on relevant encoder outputs at each generation step.

Transformer Based Approach

The Transformer architecture shown in Figure 5, introduced in the paper "Attention is All You Need" (Vaswani et al., 2017), is a non-recurrent model built for sequence-to-sequence tasks like machine translation.

Our implementation includes three main components:

- An embedding layer to convert token indices to vector representations, augmented with positional encodings
- The Transformer layers that perform multi-head attention and feedforward computation
- A linear output layer that projects final vectors into the target vocabulary space

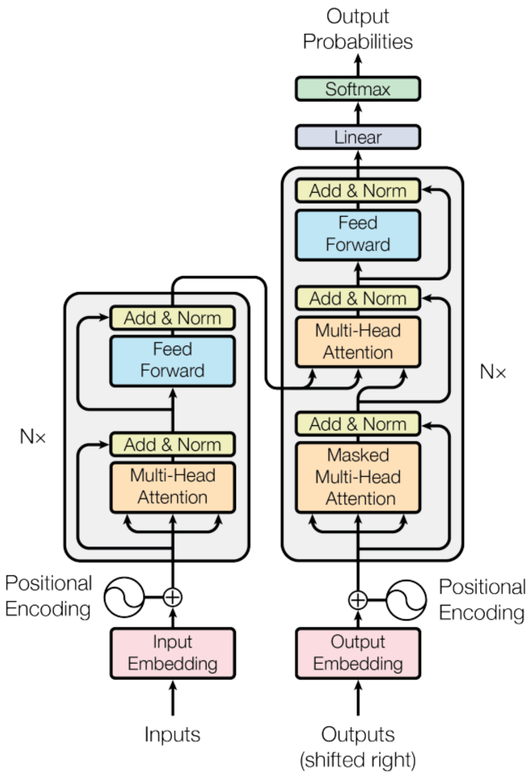

Fig. 5: Transformer Model for sequence to sequence translation (Vaswani et al., 2017)

Unlike RNN-based models, the Transformer leverages self-attention mechanisms to determine relationships between tokens in a sequence. This design allows parallelization and has proven efficient for long sequences.

The architecture uses six encoder and six decoder layers. Each encoder block consists of a multi-head self-attention mechanism and a feed forward layer. The decoder includes an additional attention layer that operates over encoder outputs. Residual connections and layer normalization are applied throughout the model for stability.

Experimental Setup

We used a pretrained LSTM encoder-decoder from PyTorch and a transformer based model from HuggingFace, fine tuning both specifically for Hindi poetry translation. Training and inference were carried out using the PyTorch framework. For the transformer setup, we employed IndicTrans2 models originally trained using Fairseq, later adapted for inference via HuggingFace Transformers (Gala et al., 2023). These models were further fine tuned on our custom Hindi poetry dataset.

For training, we used the IIT Bombay English-Hindi corpus (Kunchukuttan et al., 2018), which contains parallel English-Hindi data and monolingual Hindi texts from multiple sources, curated by the Center for Indian Language Technology, IIT Bombay.

To evaluate translation quality, we applied two common metrics: BLEU (Bilingual Evaluation Understudy) (Papineni et al., 2002) and chrF (character-level F-score) (Popović, 2015). These metrics were used to compare translation quality across different model configurations.

BLEU measures how closely the model's output matches a reference translation. It is computed as:

$$BLEU\ score = BP \times \left(\frac{1}{4} \cdot \sum_{n=1}^{4} \log p_n\right) \tag{3}$$

where BP is the brevity penalty and p_n denotes n-gram precision:

$$p_n = \frac{number\ of\ matching\ n\text{-}grams}{total\ number\ of\ generated\ n\text{-}grams} \tag{4}$$

Unlike BLEU, chrF compares character-level n-gram overlaps between machine-generated and reference translations. It is especially useful for morphologically rich languages like Hindi, where variations in word forms are common. This metric captures subtle differences by analyzing sequences at the character level.

Results

Our experiments involved multiple Hindi poems. Table 1 lists four selected poems by well-known Hindi poets, along with statistics on line and word counts.

Table 1: Details of Hindi poems considered in test

Poem#	Hindi Poet Name	#words	#lines
P1	Maithili Sharan Gupt (मैथिलीशरण गुप्त)	137	21
P2	Pt. Makhanlal Chaturvedi (माखनलाल चतुर्वेदी)	57	6
P3	Ramdhari Singh Dinkar (रामधारी सिंह दिनकर)	288	40
P4	Dwarika Prasad Maheshwari (द्वारिका प्रसाद माहेश्वरी)	121	15

Table 2 provides side-by-side comparisons between translations generated by our fine-tuned transformer model and those produced by Google Translate.

DOI: 10.3844/jcssp.2026.360.366

Table 2: A poem translation comparison with our approach and Google Translation

Source (Hindi)	Our Approach	Google Translation
"हाथ में ध्वजा रहे बाल दल सजा रहे, वीर तुम बढ़े चलो! धीर तुम बढ़े चलो!"	Children with flags in their hands. The flags never bowed. The teams never stopped. Brave you grow! Brave you grow!	Children's parties are decorating with flags in their hands, The flag never bowed, the party never stopped, Brave you move forward! You march slowly!
"सामने पहाड़ हो सिंह की दहाड़ हो तुम निडर डरो नहीं तुम निडर डटो वह वीर तुम बढ़े चलो! धीर तुम बढ़े चलो!"	The mountain in front of you is the roar of a lion, you are fearless, do not be afraid, you are fearless, you are brave, go ahead! Be patient, go ahead!	There should be a mountain in front, there should be a roar of a lion. You are fearless, don't be afraid, stand there fearless. Brave you move forward! You march slowly!
"प्रात हो कि रात हो संग हो न साथ हो सूर्य से बढ़े चलो चन्द्र से बढ़े चलो वीर तुम बढ़े चलो! धीर तुम बढ़े चलो!"	Be it morning or night, be it morning or evening, be it sun or moon, be it morning or evening, be it morning or evening!	Be it morning or night, whether you are together or not. Let's go ahead of the sun Let's go ahead of the moon. Brave you move forward! You march slowly!
"एक ध्वज लिये हुए एक प्रण किये हुए मातृ भूमि के लिये पितृ भूमि के लिये वीर तुम बढ़े चलो! धीर तुम बढ़े चलो!"	Carrying a flag, taking a vow, for the motherland, for the fatherland, O brave, march ahead! O steadfast, march ahead!	carrying a flag and taking a pledge, For motherland, for fatherland, Brave you move forward! You march slowly!
"अन्न भूमि में भरा वारि भूमि में भरा यत्न कर निकाल लो रत्न भर निकाल लो वीर तुम बढ़े चलो! धीर तुम बढ़े चलो!"	Fill up the grain in the soil, fill up the soil, try to take out the gemstone, you brave one! Be patient, you brave one!	full of food in the land, full of rain in the land, try hard and take out all the gems, Brave you move forward! You march slowly!

Figures 6 and 7 show the BLEU and chrF score comparisons, respectively. In most examples, our fine-tuned models outperformed Google's translations.

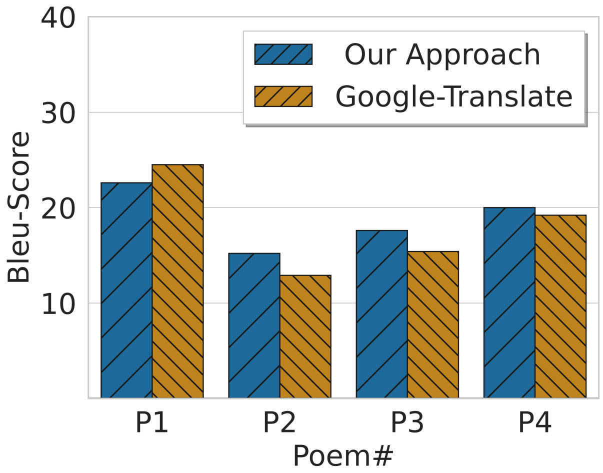

Fig. 6: Comparison of BLEU Score of Hindi poems translated to English Text

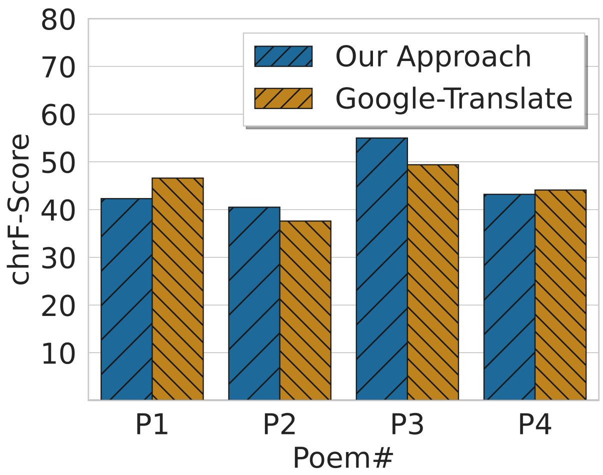

Fig. 7: Comparison of chrF Score of Hindi poems translated to English Text

Although our fine-tuned models were trained on a relatively small dataset, they consistently yielded strong results. We believe that training on a larger and more diverse poetry corpus, though computationally intensive, would further enhance performance, especially for Transformer based models.

Conclusion and Future Work

Neural Machine Translation (NMT) has made steady progress in enabling communication across languages. However, translating poetry remains one of the most difficult challenges. Hindi, spoken by millions in India, has a rich and expressive poetic tradition built over centuries. Sharing these literary works with a global audience requires translation methods that go beyond standard techniques.

This study focused on translating Hindi poetry into English using advanced NMT models. We used both RNN based encoder-decoder models with attention and transformer architectures, fine-tuning them on a custom dataset of Hindi poems. When compared to Google Translate, our models produced competitive, and often better, translations. Still, our results show the need for larger datasets and more refined models to capture poetic form and meaning more consistently.

Although NMT works well for general translation tasks, poetry demands more. A good translation should not only carry the original meaning but also reflect rhythm, metaphors, emotion, and style. Capturing these elements is a unique challenge that requires specialized approaches in computational linguistics.

Despite our promising results, some limitations remain. Our dataset was small and focused on a narrow set of poetic forms, which may reduce the model's effectiveness on other styles or dialects. Also, we relied on automated evaluation metrics like BLEU and chrF, which don't fully reflect the emotional or artistic quality of a poem. Involving human evaluators, especially literary experts, would offer a more complete picture.

Additionally, the training corpus may not represent the full diversity of Hindi poetry, which is another area for future improvement.

Poetic translation is also influenced by the reader's interpretation. Cultural references, figurative language, and historical context often carry subtle meanings that machines struggle to understand. Standard metrics also fail to account for literary richness. To address this, future research should include qualitative assessments and develop better benchmarks specific to poetry.

Potential Applications

The techniques explored in this paper have many potential uses. Educational platforms can use poetry translation to support bilingual learning and expose students to world literature. Cultural organizations and digital archives can use these tools to make classical Hindi poetry available to non Hindi speakers. Creative writing tools and multilingual reading apps can also benefit by helping users engage with poetry across languages. As NMT continues to improve, it can play a key role in bridging cultural and linguistic gaps in literature.

Recent advances in Generative AI (GenAI), such as Retrieval Augmented Generation (RAG), open new opportunities for poetry translation. These systems can add external knowledge and cultural context during translation, which may help preserve metaphors, allusions, and poetic tone. Although our current focus was on traditional NMT, future work should explore these emerging GenAI techniques to better handle the complexity of poetic language.

Acknowledgment

The authors express their sincere gratitude to the reviewers for their constructive and insightful feedback, which helped improve the quality of this manuscript. We also thank the translators and evaluators who contributed to the assessment of poetic translations. Their expertise and support were instrumental in completing this work.

Funding Information

The authors did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors Contributions

Pragya Tewari: Conceptualization; Methodology; Investigation; Formal analysis; Data curation; Writing - original draft; Visualization.

Anurag Singh Baghel: Supervision; Project administration; Writing - review & editing; Final approval of the manuscript.

References

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724-1734. https://doi.org/10.3115/v1/d14-1179

Gala, J., Chitale, P. A., Raghavan, A. K., Gumma, V., Doddapaneni, S., Kumar, A., Nawale, J., Sujatha, A., Puduppully, R., Raghavan, V., Kumar, P., Khapra, Mitesh M., Dabre, R., & Kunchukuttan, A. (2023). Indictrans2: Towards high-quality and accessible machine translation models for all 22 scheduled indian languages. *arXiv*. https://doi.org/10.48550/arXiv.2305.16307

Gangar, K., Ruparel, H., & Lele, S. (2021). Hindi to English: Transformer-Based Neural Machine Translation. *International Conference on Communication, Computing and Electronics Systems*, *733*, 337-347. https://doi.org/10.1007/978-981-33-4909-4_25

Ghazvininejad, M., Choi, Y., & Knight, K. (2018). Neural Poetry Translation. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, 67-71. https://doi.org/10.18653/v1/n18-2011

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Junczys-Dowmunt, M., Dwojak, T., & Hoang, H. (2016). Is neural machine translation ready for deployment? a case study on 30 translation directions. *arXiv*. https://doi.org/10.48550/arXiv.1610.01108

Kalchbrenner, N., & Blunsom, P. (2013). Recurrent Continuous Translation Models. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 1700-1709. https://doi.org/10.18653/v1/d13-1176

Kunchukuttan, A., Mehta, P., & Bhattacharyya, P. (2018). The IIT Bombay English-Hindi parallel corpus. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*. Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan.

Luong, T., Pham, H., & Manning, C. D. (2015). Effective Approaches to Attention-based Neural Machine Translation. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 1412-1421. https://doi.org/10.18653/v1/d15-1166

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2001). BLEU: a Method for Automatic Evaluation of Machine Translation. *In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 311-318. https://doi.org/10.3115/1073083.1073135

Popović, M. (2015). chrF: character n-gram F-score for automatic MT evaluation. *Proceedings of the Tenth Workshop on Statistical Machine Translation*, 392-395. https://doi.org/10.18653/v1/w15-3049

Seljan, S., Dunder, I., & Pavlovski, M. (2020). Human Quality Evaluation of Machine-Translated Poetry. *Proceeding of the International Convention on Information, Communication and Electronic Technology (MIPRO)*, 1040-1045. https://doi.org/10.23919/mipro48935.2020.9245436

Tewari, S., Kumar, A., & Paul, K. (2022). SACC: Split and Combine Approach to Reduce the Off-chip Memory Accesses of LSTM Accelerators. *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 580-583. https://doi.org/10.23919/date54114.2022.9774683

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Proceeding of the Advances in Neural Information Processing Systems*, 5998-6008.

Wang, H., Wu, H., He, Z., Huang, L., & Church, K. W. (2022). Progress in Machine Translation. *Engineering*, *18*, 143-153. https://doi.org/10.1016/j.eng.2021.03.023

Wang, Z., He, Wei, Wu, H., Wu, H., Li, W., Wang, H., & Chen, E. (2016). Chinese poetry generation with planning based neural network. *ArXiv:1610.09889*.

Wei, J., Zhou, Q., & Cai, Y. (2018). Poet-based Poetry Generation: Controlling Personal Style with Recurrent Neural Networks. *Proceeding of the International Conference on Computing, Networking and Communications (ICNC)*, 156-160. https://doi.org/10.1109/iccnc.2018.8390270