Efficient Fine-Tuning: Comparing LoRA and Hyperparameter Tuning for Language Translation

Asi Kuushalie
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
bl.en.u4cse22204@bl.students.amrita.edu

C V Sree Pranavi
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
bl.en.u4cse22216@bl.students.amrita.edu

Velumury Varshita
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
bl.en.u4cse22264@bl.students.amrita.edu

Peeta Basa Pati
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
bp_peeta@blr.amrita.edu

Tania Ganguly
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
BL.SC.IDCSE23002@bl.students.amrita.edu

***Abstract*——The growing size and complexity of deep learning models like MarianMT and T5 have shown good results in NLP tasks, including language translation. Those models were specifically chosen for their increased competence with complex language processing tasks: MarianMT for multi-language translations and T5 for its generic text-to-text framework. Despite their impressive performance, the fine-tuning of these models is hard due to the high computational load they carry, which can act as a substantial barrier for those with limited processing power. Techniques such as LoRA, which modifies a limited subset of model parameters for the purpose of efficient fine-tuning, can largely reduce the amount of time and memory needed for model training. The effect of incorporating the LoRA technique includes retention of the accuracy in case it exists or boosts of it.**

***Index Terms*—Parameter-efficient Fine Tuning, Model Optimization, Computational Efficiency, Selective Parameter Tuning.**

I. Introduction

Some of the powerful tools in translating text from one to many languages include language translation models like MarianMT and T5. However, their huge computational requirements pose a significant challenge when fine-tuning them for specific pairs. This makes the model hard for users to access with limited advanced computing resources. These models can't be used in resource-constrained settings because of their high computational demands. This affects localized solutions for underrepresented languages, which are very important in reducing communication gaps. Without effective ways of fine-tuning, such models do not serve the linguistic needs of diverse communities, and thus the digital divide is enhanced. It affects education, health, and e-commerce industries, which require correct translations. It also denies information to marginalized linguistic groups in their mother tongue. This would empower communities, inclusivity, and knowledge sharing across the globe. This work introduces a new approach to fine-tuning translation models by comparing standard fine-tuning methods with LoRA (Low-Rank Adaptation) fine-tuning.

The focus of the study is on optimizing MarianMT for English-French translation and T5 for Kannada-Telugu translation. LoRA fine-tuning is tested for its ability to preserve accuracy while reducing computational costs. The innovation lies in the application of LoRA fine-tuning to translation tasks with parameter-efficient optimization. What makes this paper unique is that the high-resource (English-French) and low-resource (Kannada-Telugu) language pairs are compared. This effort brings underrepresented languages as the bridge needed to close the vast gap in research translation. It increases access to educational content in the native languages, leading to contribution towards UN SDG Goal 4: Quality Education.

The paper structure is divided into: Section II covers a literature survey on the existing methods and technologies of translation models and their shortcomings for low-resource languages. Section III explains the data used in experiments with further insights into datasets and relevance of the study with them. Section IV explains how LoRA fine-tuning was applied on T5 model. Section V plots the result of Standard Fine-Tuning vs. LoRA in accuracy to better computational efficiency. Section VI discusses final observations concluded from major findings along with implications and further research directions.

II. Literature Survey

Thihlum et al. [1] researched how cost and translation performance could be balanced when merging models using decision trees and neural networks. They found some techniques, such as adapter layers and prompt tuning, which can optimize for the trade-off between good performance and computational cost. In resource-constrained scenarios, a balance between being efficient in speed and being very accurate is an important question for the practical deployment of models.The trace of large language models and their applications in chatbots by Sindhu et al. [2] were followed, as well as the ethical issues involved. Major areas are discussed with respect to model bias mitigation, fairness, and developing responsible AI. LLMs have been explored regarding how they help improve interactions with customers, including ways to deal with issues related to fairness and ethics while implementing AI.Domain-specific fine-tuning was highlighted, especially within the manufacturing domain, in Xia et al. [3]. Knowledge integration has been put into focus as a primary direction to develop more precise and context-specific outputs. In this fine-tuning method, efforts are targeted towards optimizing the LLM for specific sectors so that their outputs could be applicable to real-life situations and as practical.Nashaat and Miller [4] designed the CodeMentor framework for countering the challenges of few-shot learning and reinforcement learning with human feedback. The primary goal of this system is to counter issues like hallucination, where the model produces irrelevant or nonsensical responses, as well as the cost of training, by providing a human-guided scalable framework towards better model performance.Guo et al. [5] proposed a hybrid CNN-GRNN structure for energy consumption forecasting with anomaly detection. This approach provides the means of energy consumption more efficiently, while detecting anomalies in energy consumption, leading to better management of resources.Sidiya et al. [6] applied BERT with LSTM to improve Arabic to English translation. Their work integrated linguistic and morphological features to enhance the quality of translation, which showed that rich linguistic features can be used to overcome some of the inherent challenges in translating complex languages like Arabic.

Kumar et al. [7] applied RNNs to scalable text generation with coherence. Their work on GPT models shows how the recurrent networks can make the models generate longer texts with coherence, which is significant for applications such as content generation and chatbots.Vakayil et al. [8] provide an overview of the issues and advancements in LLMs and discuss the recent work that addresses the restrictions in these models, specifically addressing ambiguity, interpretability, and scaling models for targeted tasks.Harahus et al. [9] studied the grammatical correction capability of T5 for multiple languages. They focused on the demand for high-quality datasets as well as evaluation metrics while evaluating the performance of the model in a multilingual setup, where such models will transform the grammar correction mechanisms in other languages.Raychawdhary et al. [10] focused on XLM-R for low-resource languages, focusing on Amharic for sentiment analysis. They pointed out that XLM-R can be fine-tuned to make improvements on these languages, especially when the problem is low-resource languages, which often have a lack of training data.Pilch et al. [11] tested MarianMT for the translation of English-Polish and proposed a new evaluation metric called HUME that can better estimate the quality of translation. The authors also highlighted the challenges of gender bias faced by MarianMT in creating fair and unbiased translations.

Kosenko et al. [12] addressed privacy concerns related to large-scale GPT-4 models and proposed localized training as an approach to minimize data-sharing risks and increase model relevance to local contexts.Aji et al. [13] discussed how the tokenization affects the quality of machine translation, mainly for Indian languages, utilizing MarianMT and Seq2Seq models. They aimed to optimize the tokenization in order to improve the BLEU scores and accuracy for the translation. Lin et al. [14] Multimodal techniques combining Marian MT and CLIP4clip approach sign language translation, enhancing context understanding by including both visual input and textual input.Li et al. [15] proposed the use of T5 for biomedical text simplification. This focused on creating readable health communication through using SARI and BLEU score to make medical information highly accessible.Patel et al. [16] discussed multilingual knowledge distillation using XLM-R to enhance pre-trained sentence encoders for sentiment analysis and chatbots, which enables cross-lingual systems to be more efficient.

Yanagimoto et al. [17] have discussed the generation of textual descriptions from structured data by applying LoRA fine-tuning, which makes it useful for real-time data analysis in various industries.Fu et al. [18] proposed a framework called LoFT, LoRA-based for improving adversarial training while balancing performance with faster training, to improve generalization. Ramaneedi et al. [19] have used the mT5 model for error correction in Kannada, exemplifying how multi-lingual models may improve grammar of regional languages.Bharathi Mohan et al. [20] suggested the utilization of deep learning for the error correction in Kannada, which could improve linguistic quality in Indian language processing tools.Fuadi et al. [21] presented a lightweight version of T5 for Indonesian tasks called idT5. It presented an efficient model for low-resource settings.Suryakusuma et al. [22] presented a comparison between T5 and MBart on the task of English-to-German translation, including techniques and metrics for decoding for the enhancement of translation quality.

In summary, previous studies in neural machine translation (NMT) have laid the groundwork but often lacked efficiency and adaptability, which our paper addresses. For instance, Sidiya et al. [6] used BERT and LSTM for Arabic-English translation but did not focus on computational efficiency; our LoRA fine-tuning for T5 Kannada-Telugu translation achieves both resource optimization and quality. Similarly, Pilch et al. [11] evaluated MarianMT for English-Polish translation but struggled with reducing bias and improving metric scores, whereas our hyperparameter fine-tuning on MarianMT for English-French ensures fairness and enhanced translation accuracy. Aji et al. [13] focused on tokenization optimization for Indian languages but did not fully integrate advanced evaluation metrics like GLEU and TER. Our project not only incorporates these metrics but also combines them with fine-tuning techniques to improve low-resource language translations efficiently, demonstrating clear advantages over prior methods.

III. Data Description

The dataset used for this project consists of Kannada-to-Telugu translations, partially sourced from Kaggle. The dataset contains approximately 20,000 rows, with each entry comprising columns for Kannada text and its corresponding Telugu translation and also English to French Dataset directly sourced from Kaggle.

A. Data Collection and Preprocessing

1 Source: The Kannada dataset was obtained from Kaggle, a reputable platform for sharing datasets, ensuring a reasonable level of quality and validity.

2. Translation Process: To create the Kannada-to-Telugu dataset, the original Kannada text was translated into Telugu using the Google Translation API. This automated translation process leverages advanced neural machine translation techniques, aiming to provide accurate and contextually relevant translations.

3. Data Preprocessing:

1. Text Cleaning: The raw text data underwent preprocessing to remove extraneous characters, punctuation, or formatting issues that affect translation quality.

2. Language Consistency: Both Kannada and Telugu texts were checked for consistency and clarity to ensure that the translations were meaningful and contextually appropriate.

3. Handling Null Values: Null entries in the dataset were identified and addressed, through removal to maintain the integrity of the dataset.

IV. Methodology

This sub-section describes the fine-tuning procedure of MarianMT for English to French and the T5 model for Kannada to Telugu for translation. Subtasks include data pre-processing, utilization of pre-trained transformer models, and fine-tuning methodologies such as LoRA involved in fine-tuning of T5 models, and adoption of comprehensive metrics like GBLEU (Google-BLEU), TER(Translation Edit Rate), and accuracy. The approach of the methodology is focused in computational aspects, parameters, and control.

Figure 1 illustrates a translation pipeline for two languages in two sets of datasets, Kannada Telugu, and English French. Data acquisition is the first stage of the data preparation, then data loading, and then data transformation and cleaning. For the Kannada to Telugu dataset, the T5-small pre-trained model is employed; for the English to French dataset, the MarianMT (Helsinki) model is used. Both pipelines have the

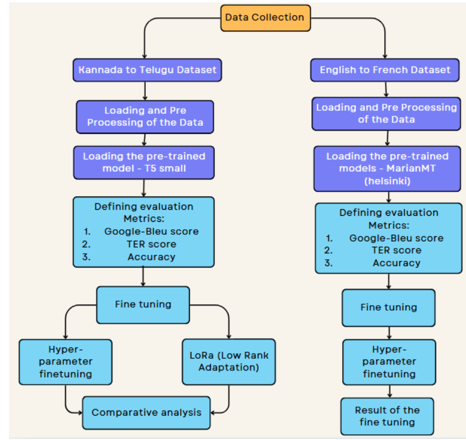


Figure 1: Flow Diagram

following evaluation parameters set Google-BLEU score, the TER score, and accuracy. The next step involves fine-tuning, hyperparameter fine-tuning and LoRA fine-tuning Kannada to Telugu and the comparison of fine-tuning strategies followed. Hyperparameter fine-tuning is performed for the English to French dataset.

This research focuses on optimizing machine translation tasks for two distinct datasets, Kannada-to-Telugu and English-to-French. For Kannada-to-Telugu translation, pre-trained transformer model T5-small is fine-tuned using techniques such as Low-Rank Adaptation (LoRA) and hyper-parameter fine-tuning, and a comparative analysis is done. For English to French language translation, the Helsinki-NLP/opus-mt-en-fr is used, however, the MarianMT is fine-tuned using hyperparameters to provide the best of the translation. To assess translations, Quality Estimation metrics, GLEU, TER, and translating Accuracy measurement is used, and hyperparameters such as learning rate and batch size, epochs are analyzed for comparison for optimized results in terms of time computation and with high translation quality in both the tasks.

Data Collection: For both the Kannada to Telugu translation dataset and English to French translation dataset, preprocessing was conducted by tokenization, converting all letters to lower cases, and removal of special symbols. Random sampling was also used to partition the dataset into training data of 80% and the testing data of 20%. This ensures the input of the model is clean and free from noise.

Tokenization: In pipeline Kannada-to-Telugu, AutoTokenizer from HuggingFace is used, to convert source and target texts into sequences, and for the equal sequence length they are padded and truncated. For English to French translation, MarianTokenizer has been used for splitting the sentence into tokens and used for text preprocessing where the maximum length is set to 128 and then followed by padding the tokenized data.

Model Loading: For Kannada to Telugu translation, T5-small is selected because of its small computational size and fine-tuning specifically for sequence-to-sequence tasks. For

English to French, the model used is Helsinki-NLP/opus-mt-en-fr, which has been made compatible with MarianMT.

Fine-Tuning: In the two pipelines, fine-tuning is done using different hyperparameters, and the results are tested and compared for effectiveness. Coming down to Kannada to Telugu, the standard and LoRA-based finetuning procedures are followed, the CPU and memory usage are also measured and the time for training both the models are compared. For the English to French translation, then hyperparameters include learning rate at 1e-5; 5e-5, batch size at 4, 8, and 16, and epochs at 2,3, and 4.

Evaluation Metrics: The assessment of the Kannada to Telugu and English to French translation pipeline also uses reliable performance for the purpose of measuring translation accuracy and computational complexity:

GLEU Score (Generalized Language Understanding Evaluation): This metric compares the n-grams from the predictions to the references while giving credit based on both precision as well as recall. It is most suitable for examining the issues of fluency and sufficient amount in the translation.

TER Score (Translation Edit Rate): TER will systematically quantify the minimum number of changes needed to transform the predictions into targets. It shows structural inconsistencies; hence it preserves syntax and semantics of the translated text.

Accuracy: To address near-matching scenarios the accuracy uses Levenshtein similarity, which goes further into providing semantic correctness while considering minor differences in the translated texts.

Results

A. T5-small model

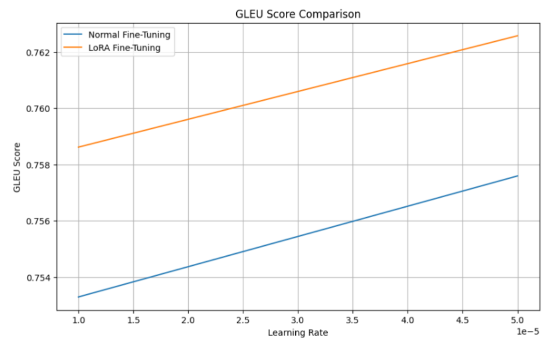
The evaluation results derived out of fine tuning experiments gives understanding on how the T5-small works for Kannada to Telugu translation, by illustrating the impact of standard fine tuning and LoRA fine tuning. Therefore, the research aims at addressing efficiency, translation quality and the compromise between the two.

Table I shows the results showing normal fine-tuning and LoRA, The former's GLEU score is between 0.753 and 0.757 and accuracy more than 94%. However, the values of Translation Edit Rate (TER) are moderate: 0,48 – 0,49 which define the presence of the structural mistakes. The use of a smaller batch size such as 4 took a lot of time (2659 seconds) more time than a larger batch size of 16, which took only 765 seconds, implying that big batch size come at a cost of time. On the other hand, the LoRA Fine-tuning Results increased the GLEUs (from 0.758 to 0.762) and reduced the TER (from 0.507 to 0.534), which supports LoRA's function of fine-tuning parameters for more effective translation quality. It is seen that LoRA retains high accuracy of 94.9% in the small batches as well and has almost equally comparable training time of 766 seconds when done in small batch size of 16.

Figure 2 also validates that the use of LoRA fine-tuning gives a satisfactory trade-off between the speed and quality of the translations. Its GLEU and TER tests make a sound testimony of the efficiency of parameter-efficient training

TABLE I: Performance Metrics for Normal Fine-Tuning and LoRA

Type	Learning Rate	Batch Size	GLEU Score	Accuracy	TER Score	Training Loss	Validation Loss
Normal	0.00001	4	0.753	0.942	0.485	0.050	0.043
Normal	0.00005	16	0.757	0.942	0.480	0.048	0.042
LoRA	0.00001	4	0.758	0.948	0.507	0.047	0.042
LoRA	0.00005	16	0.762	0.946	0.534	0.045	0.041



GLEU score comparision

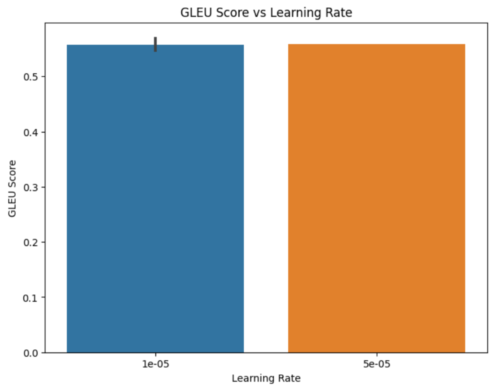
methodologies that are very relevant in resource-limited settings. The stability observed in the memory with LoRA shows the applicability of this technique in widespread machine translation.

The research reveals the role of LoRA on how it can be applied as a lightweight method to fine-tuning with the aim of enhancing both the computational strength and the quality of translation. In that way, the project partially benefits in the sense of resource-sensitive NLP solutions, especially in the cases with high demands on translation quality and low computational recourses.

B. MarianMT

The hyperparameters such as the learning rate and the batch size of the MarianMT model fine-tuned for the English to French translation task allowed the assessment of the overall impact of these two hyperparameters on the overall performance of the model.

Table II presents the results of the MarianMT fine-tuning experiments, where we showcase the role of the hyperparameters towards translation accuracy and time. Compared to the GLEU scores, our model yielded GLEU scores in the range of 0.546 – 0.569 suggesting that our model has produced appreciable linguistically valid translations although the fluctuations were observed based on learning rate and batch size. The optimal standard estimator GLEU score of the proposed model is 0.569 which requires reducing the learning rate to 0.00001 and a smaller batch size of 4, indicating that the model that is updated by very small changes achieves better performance in imitation of reference translation. Hence, observed TER scores of between 0.347 and 0.359 show that the level of structural error is low and the network's performance is marred with a fewer number of structural mistakes with enhanced results from smaller batch size due to optimization of weights on the



GLEU scores vs Learning Rate

network. Tighter alignment to the source text through fine-tuning is illustrated in Figure 1 — a bar chart of GLEU Scores and Learning rate where GLEU scores decrease slightly as the learning rate increases.

TABLE II: MarianMT Fine-Tuning Results

Learning Rate	Batch Size	GLEU Score	TER	Accuracy	Training Loss	Validation Loss
0.00001	4	0.569	0.347	0.467	0.035	0.028
0.00005	8	0.558	0.351	0.454	0.022	0.030
0.0001	16	0.547	0.360	0.443	0.014	0.031

Figure 1 highlights the bar chart of GLEU Scores and Learning rates with a slight drop in GLEU scores with higher learning rates, emphasizing the role of fine-tuning in achieving better translation alignment. The assessment of the MarianMT model proves that the model is suitable for English to French translations producing both semantically and syntactically sound equivalents of the source text. Overall, this work offers practical guidance on how hyperparameters should be set up and adjusted within a constraint because of the systematic comparison of those configurations. The use of Soft Accuracy as an additional parameter broadens the approaches like GLEU and TER, which consider the depth of translation. These results are pertinent for wide-scale MT scenarios in languages for which resources are scarce, especially those in low-resource languages.

CONCLUSION

The current study focuses on fine-tuning pre-trained transformer models to perform Kannada-to-Telugu and English-to-French translation. Key emphases are on systematic preprocessing, model adaptability, sophisticated techniques like Low-Rank Adaptation (LoRA), and optimization of hyperparameters. It proved that the transformer models have great prospects in dealing with varying language pairs, which it illustrated by using T5-small for Kannada-to-Telugu and MarianMT for English-to-French.A comprehensive evaluation by these metrics, GLEU, TER, and Accuracy, allowed evaluation of fluency, the correctness of structure, as well as semantic adequacy. The results clearly focus on the need for more balanced computational efficiency with respect to translation quality, thus opening much potential for further optimizations specifically for machine translation in general multilingual and low resource settings.

Further directions toward more sophisticated fine-tuning methods, such as utilizing adapter layers and meta-learning might further improve model adaptability and efficiency. More diverse and low-resource language datasets can be used to expand upon the robustness of the model to address nuances in their languages. Also, it can use multimodal strategy combining text with audio and visual data to understand context well for more accurate translations. Making lightweight and scalable models built for edge devices can promote real-world applications in these resource-constrained settings as well.

REFERENCES

[1] Z. Thihlum, V. Khenglawt and S. Debnath, "Machine Translation of English Language to Mizo Language," *IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*, pp. 92-97, Nov. 2020, doi: 10.1109/CCEM50674.2020.00028.

[2] B.Sindhu, RP.Prathamesh, M.B.Sameera, S.KumaraSwamy , "The Evolution of Large Language Model: Models, Applications and Challenges," *International Conference on Current Trends in Advanced Computing (ICCTAC)*,pp. 1-8, May 2024, doi: 10.1109/ICCTAC61556.2024.10581180.

[3] L.Xia, C.Li, C. Zhang, S.Liu and P. Zheng, "Leveraging error-assisted fine-tuning large language models for manufacturing excellence," Robotics and Computer-Integrated Manufacturing, vol.88, p.102728, Aug.2024.

[4] M. Nashaat and J. Miller, "Towards Efficient Fine-tuning of Language Models with Organizational Data for Automated Software Review," IEEE Transactions on Software Engineering, vol. , no. 01, pp. 1-14, Jul. 2024, doi: 10.1109/TSE.2024.3428324.

[5] J.Guo, P. Lin, L. Zhang, Y. Pan and Z. Xiao,"Dynamic adaptive encoder-decoder deep learning networks for multivariate time series forecasting of building energy consumption," Applied Energy, Vol.350, p.121803, Nov.2023.

[6] A.M. Sidiya, H. Alzaher, R. Almahdi and P.Elkafrawy, "From Analysis to Implementation: A Comprehensive Review for Advancing Arabic-English Machine Translation,"21st Learning and Technology Conference (L&T), pp. 109-114, Jan.2024 ,doi: 10.1109/LT60077.2024.10469415.

[7] P.Kumar, S.Manikandan and R.Kishore," AI-Driven Text Generation: A Novel GPT-Based Approach for Automated Content Creation," 2nd International Conference on Networking and Communications (ICNWC) ,pp. 1-6, Apr. 2024, doi: 10.1109/ICNWC60771.2024.10537562.

[8] M.A.K. Raiaan, M.S.H.Mukta, K. Fatema, N.M.Fahad, S. Sakib, M.M.J. Mim, J.Ahmad, M.E.Ali and S.Azam,"A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges," *IEEE Access*, vol. 12, pp. 26839-26874, Feb.2024, doi: 10.1109/ACCESS.2024.3365742.

[9] M. Harahus, Z.Sokolová, M.Pleva, J.Juhár, D.Hládek, J.Staš and M.Koctúrová,"Evaluation of Datasets Focused on Grammatical Error Correction Using the T5 Model in Slovak,"34th International Conference Radioelektronika (RADIOELEKTRONIKA), Zilina, Slovakia, pp.1-6, April 2024, doi: 10.1109/RADIOELEKTRONIKA61599.2024.10524071.

[10] N. Raychawdhary, A. Das, S. Bhattacharya, G. Dozier and C. D. Seals, "Enhancing Sentiment Analysis in Amharic: Leveraging Transformer-Based Language Model for Low-Resource African Languages," SoutheastCon 2024, Atlanta, GA, USA, pp. 50-55, March 2024, doi: 10.1109/SoutheastCon52093.2024.10500147.

[11] A.Pilch, R.Zygała and W.Gryncewicz,"Quality assessment of translators using deep neural networks for polish-english and english-polish translation," 12th International Conference on Advanced Computer Information Technologies (ACIT),pp. 227-230, September 2022, doi: 10.1109/ACIT54803.2022.9913189.

[12] D.P.Kosenko, Y.M.Kuratov and D.R.Zharikova,"Accessible Russian Large Language Models: Open-Source Models and Instructive Datasets for Commercial Applications," Doklady Mathematics,Vol. 108, No. Suppl 2, pp. S393-S398, December 2023.

[13] J. Saji, M. Chandran, M. Pillai, N. Suresh and R. Rajan, ”English-to-Malayalam Machine Translation Framework using Transformers,” IEEE 19th India Council International Conference (INDICON), Kochi, India, pp. 1-5,November 2022, doi: 10.1109/INDICON56171.2022.10039859.

[14] S. Lin, J. You, Z. He, H. Jia and L. Chen, ”A Novel Effective Combinatorial Framework for Sign Language Translation,”2nd International Conference on Big Data, Information and Computer Network (BDICN), Xishuangbanna, China, pp. 204-209, January 2023, doi: 10.1109/BDICN58493.2023.00050.

[15] Z. Li, S. Belkadi, N. Micheletti, L. Han, M. Shardlow and G. Nenadic, ”Investigating Large Language Models and Control Mechanisms to Improve Text Readability of Biomedical Abstracts,”IEEE 12th International Conference on Healthcare Informatics (ICHI), Orlando, FL, USA, pp. 265-274, June 2024, doi: 10.1109/ICHI61247.2024.00042.

[16] R. N. Patel, E. Burgin, H. Assem and S. Dutta, ”Efficient Multi-Lingual Sentence Classification Framework with Sentence Meta Encoders,” IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, pp. 1889-1899, December 2021, doi: 10.1109/BigData52589.2021.9671714.

[17] H. Yanagimoto, I. Kisaku and K. Hashimoto, ”Table-to-Text Using Pre-trained Large Language Model and LoRA,” 16th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), Takamatsu, Japan, pp. 91-96, July 2024, doi: 10.1109/IIAI-AAI63651.2024.00026.

[18] J. Fu, J. Fang, J. Sun, S. Zhuang, L. Geng and Y. Liu, ”LoFT: LoRA-Based Efficient and Robust Fine-Tuning Framework for Adversarial Training,” International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, pp. 1-8, June 2024, doi: 10.1109/IJCNN60899.2024.10651480.

[19] S. Ramaneedi and P. B. Pati, ”Kannada Textual Error Correction Using T5 Model,” 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, pp. 1-5, April 2023, doi: 10.1109/I2CT57861.2023.10126228.

[20] R. Elakkiya, V. Anvitha and V. Sulochana ”Fine Tuning Pretrained Transformers for Abstractive News Summarization,” International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT), Bengaluru, India, pp. 1-5, October 2023, doi: 10.1109/EASCT59475.2023.10393603.

[21] M. Fuadi, A. D. Wibawa and S. Sumpeno, ”Adaptation of Multilingual T5 Transformer for Indonesian Language,” 9th Information Technology International Seminar (ITIS), Batu Malang, Indonesia, pp. 1-6, October 2023, doi: 10.1109/ITIS59651.2023.10420049.

[22] M. R. Suryakusuma, M. Faqih Ash Shiddiq, H. Lucky and I. A. Iswanto, ”Investigating T5 Generation Neural Machine Translation Performance on English to German,” International Conference on Informatics, Multimedia, Cyber and Informations System (ICIMCIS), Jakarta Selatan, Indonesia, pp. 12-15, November 2023, doi: 10.1109/ICIMCIS60089.2023.10349061.

[23] D.Phogat, K.S.Prashanth, M.S.Rishith, R.C.Sai, S.B.Karthikeya, G.Jyothish Lal and B.Premjith, ”Bridging Language Barriers: Exploring Hindi-to-English Speech-to-Speech Translation for Multilingual Communication,”In Congress on Intelligent Systems Singapore: Springer Nature Singapore, pp. 141-152, September 2023.

[24] S.R.Laskar, B.Paul, P.Dadure, R.Manna, P.Pakray and S. Bandyopadhyay,”English–Assamese neural machine translation using prior alignment and pre-trained language model,”Computer Speech & Language, vol.82, p.101524, July 2023.

[25] I. V. Srisurya and R. Prasanna Kumar, ”Neural Machine Translation using Adam Optimised Generative Adversarial Network,” 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, pp. 383-387, Feb. 2023, doi: 10.1109/ICCMC56507.2023.10084034.

[26] R.Jairam, G.Jyothish and B.Premjith,”A Few-Shot Multi-Accented Speech Classification for Indian Languages using Transformers and LLM's Fine-Tuning Approaches,” Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pp. 1-9, March 2024.