International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3689-3695
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0482

Few-Shot Prompt Engineering for Multi Language Text Classification Using LlM

Dr M Prabu[1], Tharun M A[2], Karthikeyan M[3], Kalai M[4]
[1] Associate professor, Department. of CSE, SRM Institute of Engineering & Tech., Chennai, Tamil Nadu, India
[2, 3, 4]UG Scholar, Department of CSE, SRM Institute of Engineering & Tech., Chennai, Tamil Nadu, India

Email ID: *prabum@srmist.edu.in[1], tm4064@srmist.edu.in[2], km1858@srmist.edu.in[3], km4414@srmist.edu.in[4]*

Abstract
Text classification is one of the important tasks in natural language processing which plays a main role in applications like sentiment analysis, spam detection, document categorization and multi-information retrieval. Traditional machine learning models require large labeled data and more feature engineering to get best accuracy. However, many real-world applications include low resource language where labeled data are scarce. This research a system for muti language text classification using few shot prompts engineering with large language models. The proposed system leverages the capability of modern large language models to understand contextual language patterns and perform classification tasks using small examples. Instead of training huge neural networks from starting, carefully designed prompts which has small number of labelled examples guide the model to perform this task effectively. The system supports multiple languages, classification text written in all the languages. An interactive interface allows the users to give multilingual text, which is processed through a prompt-based classification pipeline. The model analyzes contextual meaning, linguistic patterns, and semantic relationship to find the correct category. Experimental results shows that this few-shot prompt engineering significantly reduces the use of large data while maintain reliable accuracy.
Keywords: *Few-Shot Learning, Prompt Engineering, Large Language Models, Multilingual Text Classification, Natural Language Processin*

1. Introduction

The rapid growth of digital communication has result in huge increase in text data generation across multi languages[1]. Social media platforms, online forums, news, websites and digital service produce large content of text that needs processing and analysis[2]. Text classification is one of the most needed tasks in natural language processing, enabling systems to category text into predefined labels such as topics and intent. Traditional text classification methods rely heavily on supervised learning Figure 1 . Algorithms such as support vector, decision trees and deep learning architectures needs more labelled data for their training itself Figure 3. While these methods achieve strong performance when sufficient data is available, they become less effective in scenarios where labelled data is limited or unavailable Figure 4. So, this few shot prompt engineering can be used in multilingual languages and also to languages which has very low dataset Figure 4[3]. Modern technology has created new AI called large language models which is the brain of the proposed model which can understand the complex structure of the sentence and relationship between the text in the sentence which is used to understand the context of the sentence. These models are trained using huge multilingual data which has different languages and it can perform a wide range of natural language processing tasks Figure 5. Instead of requiring specific training for each language, this large language model can perform through prompt engineering, where instructions and examples are embedded directly in the input prompt given by the user[4]. Few shot prompt engineering has become efficient technique leveraging large language models with minimal training. In few shot prompt engineering, models are given with less example within the prompt to guide prediction[5]. These examples are used to find patterns and perform classifications on new input[6]. The integration of the few shot model with the text process opens a new possibility for scalable and natural language applications in the real world[7]. Places like India the users communicate in many different languages and accents[8]. Developing systems which is capable of understanding and classify multiple languages is needed for improving digital accessibility and information management[9]. This research provides a few shot prompt engineering frameworks for multi-language text classification. The system gives a context understanding capability of the different language model to classify using very small labelled data[10]. By giving structured prompts that has instruction and mapping, this system can give accurate category across different languages. Furthermore, multi-language system must address syntax variations and interpretation differences across multiple languages[11]. Languages such as Tamil, Hindi and English has different grammatical structures, variations and vocabularies differences[12]. Large language models will address these challenges faced in the traditional method by learning all languages during the training period[13]. These representations take meanings, syntactic relationships and context dependencies of the sentence across multiple languages. As a result, LLMs can process multi language inputs without requiring different models for each of the universal languages available[14]. Another important advantage of the prompt-based systems is the adaption. Unlike old traditional machine learning patterns that needs training for new and each task but prompt engineering allows the users to modify task behavior simply by changing the prompt given in the input by the users. This makes the model flexible and suitable for rapid prototyping and dynamic tasks[15][16].

2. Literature Survey

Text classification has been extensively studies in the natural language processing fields. Old research relied on the statistical ML learning models such as Nive Bayes, logistic regression and support vector. These models need manual feature extraction techniques such as bag of words and TF-IDF representations to convert text into numerical features[17]. With recent advancements of deep learning and neural network architecture such as recurrent neural networks and convolutional neural networks were introduced for this. These models help few shot prompt model to improve performance by learning features automatically from text sequences. However, deep learning models typically require large data and significant computational resources for training[18]. Recent developments in transformers have shown advanced NLP capabilities[19]. Transformers models utilize attention mechanisms to capture contextual relationships between words[20]. This architecture allows the model to process whole text at the same time and learn deeper semantic representations. Large language models built upon transformers architectures have shown great performances across range of NLP tasks. These models are pre-trained on multi datasets and can perform tasks like translation, summarization, question answering and text classification with minimal task-specific training. Few-shot learning has emerged as an efficient approach to utilize large language models. Instead of training models on large datasets, few-shot methods provide little numbers of examples within prompt to guide the models' predictions. Research studies have shown that carefully designed prompts can improve the classification performance with limit data. Prompt engineering has also explored in multilingual contexts, where models are tasked with processing text from multiple languages. Studies indicate that classification tasks while minimizing the need for extensive labeled data. However, most existing research focuses on English-language data or high-resource languages. These remains a need for frameworks that support multilingual classification tasks while minimizing the need for extensive datasets. This research addresses this challenge by designing a prompt-based classification system capable of handling inputs with less examples. Hybrid system outperforms than the standalone model depending on how well the fusion is performed.

3. Proposed Methodology

This section outlines the methodology used for Few-Shot Prompt Engineering based Multi-Language

Text Classification system using large Language Models (LLM)[21].

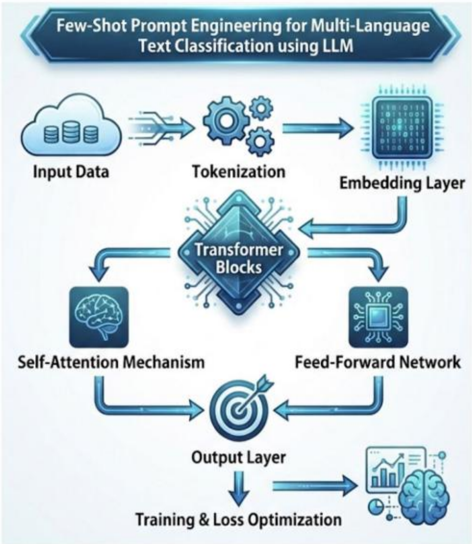

Figure 1 Architecture of Few-Shot Prompt Engineering Based Multi-Language Text Classification System using LLM

3.1. Text input and collection unit

This text input module is responsible for acquiring multilingual text data from users or external sources. This model support text written in multiple languages such as English, Tamil, Hindi and other languages. Text data originate from various platforms including user query, social media posts, online reviews, customer feedback and documents. This system accepts raw textual input through a user interface or API[22].

3.2. Text Preprocessing and Normalization Unit

The preprocessing module performs essential cleaning to improve the quality and quantity of data. Text preprocessing ensures that the language model receives standardization input for processing. This operation includes removal of unnecessary punctuation, whitespace, text formatting and encoding. These helps remove noise.

3.3. Tokenization and Linguistic Segmentation

Tokenization converts the input text into smaller semantic units known as tokens. Tokens may represent words, subworlds, or characters depending on the tokenization strategy used by the language model. Modern LLM like byte pair encoding (BPE) or word piece tokenization. These approaches allow the system to efficiently process rare words and vocabulary.

3.4. Embedding Representation Layer

The embedding layer converts tokenized words into dense vector representations that will take the semantic meaning and relationships between the words in the sentence. Words embeddings represent each other token as multi-dimensional in semantic space. Words with similar meaning tend to have similar vector representations. This makes the model to understand between the words even in different languages.

3.5. Transformer-Based Language Processing Unit

The transformer architecture forms the core computational engine of the system. They use attention mechanism to analyze the tokens and give the representational of the input text. capture both local and global dependencies within the text which enables the model to understand the context of the sentence.

3.5.1. Self-Attention Mechanism for Contextual Understanding

The self-attention mechanism is the key of transformers. It enables the model to determine the relative importance of each word in relation to the other word in the sentence. They assign weights to token based of the importance of the word in that sentence. More the weight of the token more the value of the token. This helps the self-attention mechanism to enable the system to capture the long-distance relationships between words in the sentence and understand the complex semantic patterns within the multilingual text. This mechanism plays a very crucial role in increasing the accuracy of classification by making the model to analyze context dependencies across the entire input sentence.

3.5.2. Feed-Forward Neural Processing Layer

After the self-attention mechanism is used to make the token representation, it is now passed to the feed-forward neural networks. These networks use nonlinear transformations that refine the context features extracted from the input given by the users. The feed-forwarding layers consist of neural networks which are fully constructed that will help to transform the attention outputs into a higher-level semantic representation. This stage improves the model's ability to detect and correctly find the patterns which are associated with multiple text categories. By combining the attention mechanism from the self-attention mechanism and the nonlinear transformers from feed-forwarding neural networks it will enhance the model's capability to perform a really high-level text classification than the traditional methods[23].

3.5.3. Few-Shot Prompt Engineering Module

Few-shot prompt engineering is the main central system of the proposed system. Instead of building a new model and training that model completely and separately for each particular task for classification, this system leverages and gives a pre-trained large language model and tell the LLM what to do by giving prompts. This prompt which are given to the LLM contains examples which are labeled that demonstrate the relationship between input and text classification pattern without requiring huge training and data. This deep architecture demonstrates strong performance in identifying visually complex temple structures and distinguishing similar architectural styles across different dynastic influences[24].

3.6 Classification Output Generation

After processing the prompt and contextual text representations, this model will generate a final classifies output. The system assigns the most appropriate category label to the text which is given in the input based on the semantic understanding of the sentence by the model. The classification output includes sentiment categories which contains positive, negative or neural, or topic-based categories such as technology, education, sports or politics.

3.7 Model Evaluation and Performance Metrics

The performance of this proposed text classification system is evaluated using several statistical metrics and accurate. The effectiveness of the Few-shot prompt engineering for text classification using large language model was evaluated using multiple statistical performance metrics to ensure reliable model assessment.

Classification Accuracy: Accuracy measures the overall correctness of the few-shot prompt engineering for classification by calculating the proportion of correctly classified samples relative to the total number of samples available in the dataset.

Precision: Precision evaluates the correctness of the proportion of correctness of classification predictions by calculating the proportion of the correctly classified sample by the few-shot prompt engineering for classification given by the total number of samples.

Recall: Recall measures the ability of the model to correctly identify actual few-shot prompt engineering for classification present in the dataset.

F1-Score: F1-score combines the precision and recall to find the balanced evaluation of the classification performance. This can be used particularly when dealing with imbalance in class.

Confusion Matrix Analysis: Confusion matrix helps identify misclassification trends, especially between temples with visually similar architectural features.

Inference Time and Model Efficiency: Model execution time and computational resource consumption were evaluated to determine deployment feasibility for real-time few-shot prompt engineering for classification applications.

Thus, methodology integrates processing, deep learning classification, and chatbot-based information retrieval to enable efficient few-shot prompt engineering for classification.

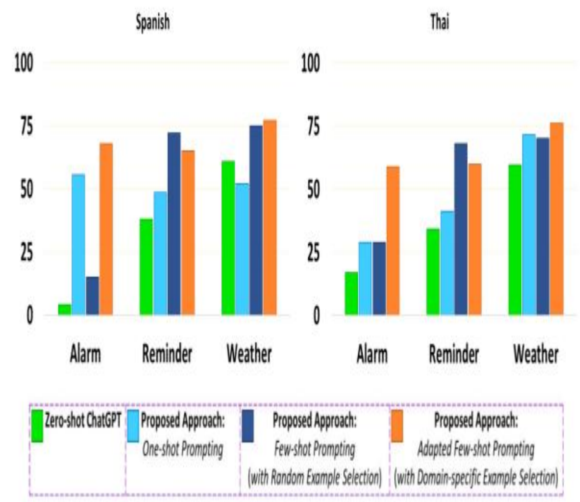

Figure 2 score comparison across domains and prompting strategies in the proposed approach.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3689-3695
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0482

4. Results And Discussion

This proposed few-shot prompt engineering system was checked using multilingual text data which contains different language samples and its categories. We have taken many experiments to find the accuracy, prompt effectiveness and the processing capability of the system[25].

4.1.Model Performance

The results of the experiments taken from the few shot prompt engineering shows that this model have achieved very reliable classification performance even within a small amount of data given to it. This language model has successfully interpreted context of the sentence and produced accurate predictions for both English and regional languages all over the world. We have taken experiments from different models also and compared it with this few shot prompt engineering model and we found out that this prompt design has significantly increased the classification performance. The prompts which tell the correct descriptions of the query and provide correct available examples to perform the task have more accurate predictions. We can also improve the reliability for classification task by increasing the number of examples within the prompt. After doing multilingual evaluation we have come to know that few shot prompt engineering model have effectively processed the written text in different languages without needing large data of language specific training. Due to the systems multilingual knowledge, the system has shown strong cross-lingual capabilities. In terms of efficiency, we reduced the extensive training of the model by giving prompt-based approach[26]. Instead of creating a separate model for each language category using large number of datasets which also takes way more time for classification, the same language model can be reused within different prompts.

4.2.Output Prediction

The output prediction has shown the proposed few shot prompt engineering model have strong understanding of the context of the given sentence whatever the language it maybe[29]. This large language model can able to find the semantic relationships within the given sentence and find the most appropriate category of classification based on the few shot examples provided in the prompt by the user[27]. By doing various experiments and observing them we have come to know that the model has effectively generalized from the limited example given by the user and produce accurate predictions even for the previous unseen text samples. This feature has highlighted ability of the use of large language model to perform this classification task with minimal training data by leveraging their pre-trained knowledge and context reading capabilities. Another significant observation taken from the experiment done in using the few shot prompt engineering was to check the system's ability to handle linguistic diversity in different and multiple languages. The model will now successfully process the text input given in English, Tamil, Hindi and any other language. Because of the multilingual ability in the large language model, it can capture the semantic meaning across every language and can generate classification outputs consistently. However, we can also find very minor inconstancies particularly when the sentence contains mixed languages in a same sentence or informal expressions. But we can fix these issues by improving the design's prompt and give more few shot examples as the input. In addition, this prompt engineering role played a huge role in giving the classification reliability. We can improve the prediction and accuracy of the few shot prompt engineering model by carefully structuring the prompt containing with clear instruction of the task and very well-defined examples of pairs. The experiments have revealed that the increasing number of examples in the prompt have significantly improved the consistency of the proposed model. The output prediction confirms that the integration of prompt engineering with large language models provides efficient and scalable approach for multilingual text classification tasks[30].

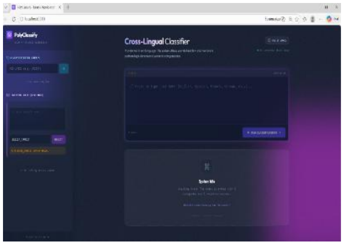

Figure 3 User Interface of the Cross-Lingual Classification System

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3689-3695
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0482

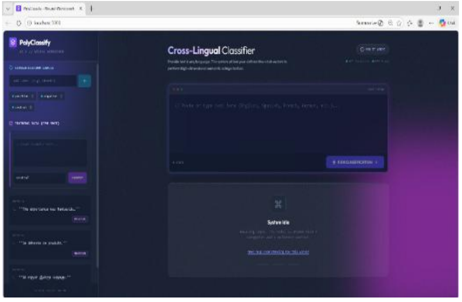

Figure 4 Label Creation and Few-Shot Training Process

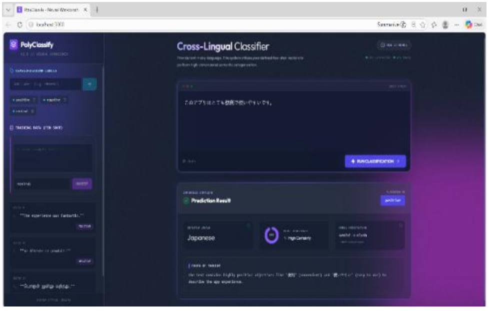

Figure 5 Cross-Lingual Classification and Prediction Output

Conclusion

This research presented a few shot prompt engineering frameworks for multilingual text classification using large language models. The proposed system has shown how the prompt-based learning can perform classification tasks effectively without requiring large labelled datasets. The evaluation of the experiment confirms that few shot promptings can achieve reliable classification performance while maintaining the efficiency. The system also supports multi language interactions and improving the accessibility for the diverse user groups. Future research may focus on the improving of the prompt given to the model which expands the language coverage and integrating the framework into real time multilingual applications.

Acknowledgment

We would like to express our gratitude to our project guide, Dr. M. Prabu, for his constant support, helpful guidance, and direction throughout the duration of this project. His guidance kept us on track and enabled us to finish our work successfully. We also thank the faculty and fellow students at SRM IST, Ramapuram, for their useful suggestions and support whenever required. A special acknowledgment goes to our families and friends for their support and encouragement. We would finally appreciate the open-source community for supplying the tools and resources used so that we are able to make this research work[28].

References

[1]. P. Pakray, A. Gelbukh, and S. Bandyopadhyay, ‘‘Natural language processing applications for low-resource languages,’’ Natural Lang. Process., vol. 31, no. 2, pp. 183–197, Mar. 2025.

[2]. S. Y. Kwon, G. Bhatia, E. M. B. Nagoudi, A. A. Inciarte, and M. Abdul-Mageed, ‘‘Zero-shot slot and intent detection in low-resource languages,’’ in Proc. 10th Workshop NLP Similar Lang., Varieties Dialects (VarDial), Jan. 2023, pp. 241–250.

[3]. L. Tu, J. Qu, S. Yavuz, S. Joty, W. Liu, C. Xiong, and Y. Zhou, ‘‘Efficiently aligned cross-lingual transfer learning for conversational tasks using prompt-tuning,’’ in Proc. Findings Assoc. Comput. Linguistics, EACL, 2024, pp. 1–21.

[4]. W. Tufa, I. Markov, and P. Vossen, ‘‘Unknown script: Impact of script on cross-lingual transfer,’’ in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics: Human Lang. Technol., 2024, pp. 124–129.

[5]. O. Xhelili, Y. Liu, and H. Schuetze, ‘‘Breaking the script barrier in multilingual pre-trained language models with transliteration-based post-training alignment,’’ in Proc. Findings Assoc. Comput. Linguistics, EMNLP, 2024, pp. 11283–11296

[6]. S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. Tat Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. Tulio Ribeiro, and Y. Zhang, ‘‘Sparks of artificial general intelligence: Early experiments with GPT-4,’’ 2023, arXiv:2303.12712.

[7]. X. V. Lin et al., ‘‘Few-shot learning with multilingual generative language models,’’ in Proc. Conf. Empirical Methods Natural Lang. Process., 2022, pp. 9019–9052.

[8]. K. Goswami, L. Lange, J. Araki, and H. Adel, ‘‘SwitchPrompt: Learning domain-specific gated soft prompts for classification in low-resource domains,’’ in Proc. 17th Conf. Eur. Chapter Assoc. Comput. Linguistics, 2023, pp. 2689–2695.

[9]. S. Schuster, S. Gupta, R. Shah, and M. Lewis, ‘‘Cross-lingual transfer learning for multilingual task-oriented dialog,’’ in Proc. Conf. North, 2019, pp. 3795–3805.

[10]. M. Akbari, A. H. Karimi, T. Saeedi, Z. Saeidi, K.

Ghezelbash, F. Shamsezat, M. Akbari, and A. Mohades, ‘‘A persian benchmark for joint intent detection and slot filling,’’ 2023, arXiv:2303.00408..

[11]. J. Pei, G. Yan, M. De Rijke, and P. Ren, ‘‘Mixture-of-Languages routing for multilingual dialogues,’’ ACM Trans. Inf. Syst., vol. 42, no. 6, pp. 1–33, Nov. 2024.

[12]. M. Calvo, L.-F. Hurtado, F. Garcia, E. Sanchis, and E. Segarra, ‘‘Multilingual spoken language understanding using graphs and multiple translations,’’ Comput. Speech Lang., vol. 38, pp. 86–103, Jul. 2016.

[13]. M. Calvo, F. García, L.-F. Hurtado, S. Jiménez, and E. Sanchís, ‘‘Exploiting multiple hypotheses for multilingual spoken language understanding,’’ in Proc. 17th Conf. Comput. Natural Lang. Learn., Aug. 2013, pp. 193–201.

[14]. L. Xiang, J. Zhu, Y. Zhao, Y. Zhou, and C. Zong, ‘‘Robust cross-lingual task-oriented dialogue,’’ ACM Trans. Asian Low-Resource Lang. Inf. Process., vol. 20, no. 6, pp. 1–24, Nov. 2021.

[15]. Z. Liu, J. Shin, Y. Xu, G. I. Winata, P. Xu, A. Madotto, and P. Fung, ‘‘Zeroshot cross-lingual dialogue systems with transferable latent variables,’’ in Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP), 2019, pp. 1297–1303.

[16]. Z. Liu, G. I. Winata, P. Xu, Z. Lin, and P. Fung, ‘‘Cross-lingual spoken language understanding with regularized representation alignment,’’ in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), 2020, pp. 7241–7251.

[17]. L. Qin, M. Ni, Y. Zhang, and W. Che, ‘‘CoSDA-ML: Multi-lingual codeswitching data augmentation for zero-shot cross-lingual NLP,’’ in Proc. Twenty-Ninth Int. Joint Conf. Artif. Intell., Jul. 2020, pp. 3853–3860.

[18]. P. Nguyen Van, T. Cao Hoang, D. Nguyen Manh, Q. Nguyen Minh, and L. Tran Quoc, ‘‘ViWOZ: A multi-domain task-oriented dialogue systems dataset for low-resource language,’’ 2022, arXiv:2203.07742.

[19]. B. Ding, J. Hu, L. Bing, M. Aljunied, S. Joty, L. Si, and C. Miao, ‘‘GlobalWoZ: Globalizing MultiWoZ to develop multilingual task-oriented dialogue systems,’’ in Proc. 60th Annu. Meeting Assoc. Comput. Linguistics, 2022, pp. 1639–1657.

[20]. J. Devlin, M. Chang, K. Lee, and K. Toutanova, ‘‘BERT: Pre-training of deep bidirectional transformers for language understanding,’’ in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Jan. 2018, pp. 4171–4186.

[21]. H. Li, A. Arora, S. Chen, A. Gupta, S. Gupta, and Y. Mehdad, ‘‘MTOP: A comprehensive multilingual task-oriented semantic parsing benchmark,’’ in Proc. 16th Conf. Eur. Chapter Assoc. Comput. Linguistics, 2021, pp. 2950–2962.

[22]. H. Zhou, I. Iacobacci, and P. Minervini, ‘‘XQA-DST: Multi-domain and multi-lingual dialogue state tracking,’’ in Proc. Findings Assoc. Comput. Linguistics, 2023, pp. 999–1009.

[23]. C.-C. Hung, A. Lauscher, I. Vulić, S. Ponzetto, and G. Glavaš, ‘‘Multi2WOZ: A robust multilingual dataset and conversational pretraining for task-oriented dialog,’’ in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol., 2022, pp. 3687–3703.

[24]. Z. Lin, A. Madotto, G. Indra Winata, P. Xu, F. Jiang, Y. Hu, C. Shi, and P. Fung, ‘‘BiToD: A bilingual multi-domain dataset for task-oriented dialogue modeling,’’ 2021, arXiv:2106.02787.

[25]. Y. Liu, J. Gu, N. Goyal, X. Li, S. Edunov, M. Ghazvininejad, M. Lewis, and L. Zettlemoyer, ‘‘Multilingual denoising pre-training for neural machine translation,’’ Trans. Assoc. Comput. Linguistics, vol. 8, pp. 726–742, Dec. 2020.

[26]. S. Tahery and S. Farzi, ‘‘An invasive embedding model in favor of lowresource languages understanding,’’ Faculty Comput. Eng., K. N. Toosi Univ. Technol., 2025. [Online]. Available: https://github.com/saedeht/ Cross-Lingual-NLU_Invasive-Embedding-Model.

[27]. L. Zuo, K. Qian, B. Yang, and Z. Yu, ‘‘AllWOZ: Towards multilingual task-oriented dialog systems for all,’’ 2021, arXiv:2112.08333.

[28]. L. Xue, N. Constant, A. Roberts, M. Kale, R. Al-Rfou, A. Siddhant, A. Barua, and C. Raffel, ‘‘MT5: A massively multilingual pre-trained Textto-Text transformer,’’ in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol., 2021, pp. 483–498.

[29]. K. Yu, H. Li, and B. Oguz, ‘‘Multilingual Seq2seq training with similarity loss for cross-lingual document classification,’’ in Proc. 3rd Workshop Represent. Learn. NLP, 2018, pp. 175–179.

[30]. B. McCann, J. Bradbury, C. Xiong, and R. Socher, ‘‘Learned in translation: Contextualized word vectors,’’ Adv. Neural Inf. Process. Syst., vol. 30, pp. 6297–6308, Dec. 2017.