International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Analyzing Multilingual LLMs Using Pre-Trained Dataset Model

Hitesh Kumar Balegar [1], Vinayak S [2]

Department: Information Science and Engineering
Institute: Vishveshwaraya Technological University
College: JSS Academy of Technical Education, Bangalore
Email : gopinathhitesh@gmail.com , vinayakshivakumar8@gmail.com

ABSTRACT:

This paper presents a comprehensive analysis of Multilingual Language Models (LLMs) utilizing a pre-trained dataset model. Multilingual LLMs have gained prominence in various natural language processing tasks, demonstrating their efficacy in handling multiple languages. In this study, we investigate the performance of these models across diverse languages and examine their generalization capabilities.

We leverage a pre-trained dataset model to evaluate the multilingual LLMs, providing insights into their behavior on various linguistic tasks. The analysis encompasses aspects such as language-specific nuances, cross-lingual transfer learning, and the impact of training data size on model performance.

Our findings reveal the strengths and limitations of multilingual LLMs, shedding light on their adaptability to different language families and their potential for practical applications in multilingual environments. This research contributes to a better understanding of the dynamics of these models and offers valuable guidance for leveraging them effectively in multilingual natural language processing tasks.

Keywords: Multilingual Language Models, Pre-Trained Models, Natural Language Processing, Cross-Lingual Transfer Learning, Linguistic Analysis.

1. Introduction

The proliferation of multilingual language models (LLMs) has revolutionized natural language processing (NLP) by enabling powerful cross-lingual capabilities. In this context, this paper aims to comprehensively analyze the effectiveness of pre-trained dataset models in the context of multilingual LLMs.

This study investigates the impact of pre-training data on the performance of multilingual language models across various languages. By examining the characteristics of diverse datasets and their influence on model behavior, we aim to provide insights into the nuances of cross-lingual learning and representation within LLMs.

Our research delves into the intricacies of multilingual language processing, exploring the challenges and opportunities that arise in the development and utilization of pre-trained dataset models. Through a systematic analysis of various benchmark tasks, including but not limited to machine translation, language understanding, and text generation, we aim to shed light on the capabilities and limitations of multilingual LLMs under the influence of pre-trained dataset models.

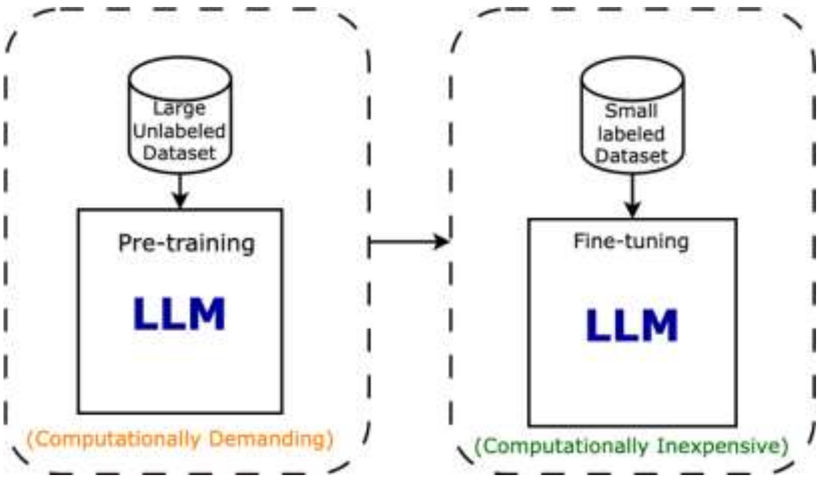

Figure 1: Large Language Models in Deep Learning

The findings presented in this paper contribute to a deeper understanding of the underlying mechanisms governing the performance of multilingual LLMs, offering valuable implications for the advancement of cross-lingual NLP applications and the refinement of pre-training strategies. Furthermore, the insights garnered from this study can serve as a roadmap for the development of more robust and efficient multilingual language models, facilitating enhanced communication and comprehension across diverse linguistic landscapes.

Some key areas of research and development within this field include:

1. Cross-Lingual Transfer Learning: Techniques that enable LLMs to transfer knowledge and linguistic patterns learned in one language to another, making it possible to improve model performance for low-resource languages.

2. Multilingual Fine-Tuning: The process of fine-tuning LLMs on specific languages to enhance their language-specific performance.

3. Cross-Lingual Document Retrieval: Developing models that can retrieve documents written in different languages to provide users with information in their preferred language.

4. Multilingual Chatbots: Building conversational agents capable of understanding and generating text in multiple languages, facilitating global customer support and user interaction.

5. Cultural Sensitivity: Ensuring that LLMs produce content that respects the cultural nuances, context, and norms of different language communities.

2. Research Problem and Main Objectives:

2.1 Research Problem:

The research problem in this context could be to understand, evaluate, and improve the performance and capabilities of Multilingual Language Models in various natural language processing (NLP) tasks, such as translation, sentiment analysis, question answering, and text generation. This problem arises due to the growing importance of LLMs in a multilingual and diverse world, where they play a critical role in bridging language barriers and enabling communication and information access across languages.

2.2 Main Objectives:

The specific objectives and focus of the research will depend on the context, available resources, and the goals of the researchers. Analyzing Multilingual LLMs is a broad and evolving field with diverse applications, and researchers can tailor their objectives to address specific challenges and opportunities in this domain.

The main objectives of a research project analyzing Multilingual LLMs using pre-trained dataset models may include the following:

1. Assess the performance of various Multilingual Language Models in different NLP tasks across multiple languages. This involves measuring their accuracy, efficiency, and generalizability.
2. Compare different LLMs (e.g., BERT, GPT, RoBERTa, mBERT) to identify their strengths and weaknesses in handling multilingual data and tasks.
3. Investigate how well LLMs can transfer knowledge from one language to another and the factors influencing this transfer. Explore techniques for improving cross-lingual transfer learning.

4. Explore techniques for fine-tuning LLMs on specific tasks or domains in multilingual settings. Assess the benefits of fine-tuned models in achieving task-specific performance.
5. Examine potential biases, fairness, and ethical concerns in using LLMs for multilingual applications. Develop strategies to mitigate bias and promote fairness.
6. Address the challenges of using LLMs in languages with limited linguistic resources. Investigate methods for improving the performance of LLMs in low-resource languages.
7. Contribute to the development of pre-trained datasets tailored for specific multilingual applications, considering the linguistic and cultural diversity.
8. Investigate methods to make LLMs more interpretable and provide explanations for their predictions and decisions, which is especially important for sensitive applications.
9. Explore user-friendly interfaces and tools that facilitate the use of LLMs for multilingual communication, making them accessible to a wider audience.
10. Apply the insights gained from the research to real-world multilingual applications, such as machine translation, content recommendation, and language understanding in various industries, including healthcare, legal, and education.
11. Investigate ways to make LLMs more scalable and resource-efficient to handle large-scale multilingual tasks effectively.

Figure 2: Unleashing the Power of Large Language Models

3. Analysis of multilingual LLMs

Overall, the analysis of multilingual LLMs using pre-trained dataset models plays a crucial role in advancing the capabilities of language models, promoting cross-linguistic understanding, and fostering inclusivity and accessibility in the digital space.

Analyzing multilingual large language models (LLMs) using pre-trained dataset models holds significant importance in several aspects:

1. Multilingual LLMs are designed to understand and generate text in multiple languages. Analyzing these models with pre-trained datasets enables researchers and developers to evaluate their capabilities in various languages, helping to improve language understanding and generation tasks across different linguistic contexts.
2. Analyzing multilingual LLMs with pre-trained dataset models facilitates cross-linguistic transfer learning, where knowledge gained from one language can be applied to another. This enables the effective transfer of linguistic patterns and structures between languages, enhancing the model's performance and efficiency in handling diverse linguistic tasks.
3. Multilingual LLMs offer the advantage of catering to a diverse range of languages and cultures. Analyzing these models using pre-trained dataset models allows for a better understanding of how well the LLMs capture the nuances and intricacies of various languages, thereby ensuring inclusivity and sensitivity to different cultural contexts.

4. By analyzing multilingual LLMs using pre-trained dataset models, researchers can fine-tune the models for specific languages, improving their performance in tasks specific to those languages. This fine-tuning process can enhance the accuracy and fluency of the model's output, making it more effective for language-specific applications and tasks.
5. Multilingual LLMs have the potential to break language barriers, enabling global communication and facilitating access to information in different languages. Analyzing these models with pre-trained dataset models contributes to the development of more accurate and comprehensive translation, interpretation, and cross-lingual information retrieval systems, making information more accessible to a broader audience.
6. By analyzing multilingual LLMs using pre-trained dataset models, researchers can enhance various NLP applications, including machine translation, sentiment analysis, text summarization, and question-answering systems, among others. This leads to the development of more robust and effective NLP applications that can cater to diverse linguistic needs and requirements.

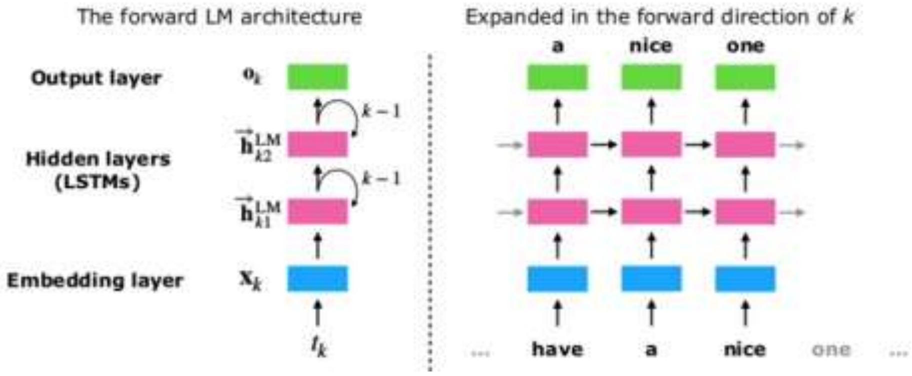

Figure 3: Introducing Poly LM: An Open Source Multilingual LLM for Diverse Language

4. Review of Research Area

Multilingual Language Models (LLMs) have gained significant attention in the field of Natural Language Processing (NLP) due to their ability to understand and generate text in multiple languages. These models, like mBERT, XLM-R, mT5, and mBART, have been developed to handle diverse languages and tasks. Here's a review of the existing literature on multilingual LLMs and their applications in NLP.

1. Multilingual LLMs are often pretrained on large multilingual corpora, which allows them to learn shared linguistic features across multiple languages.
2. Literature has explored different techniques to optimize multilingual pretraining, such as training on parallel data or using language-agnostic objectives.
3. One of the primary applications of multilingual LLMs is cross-lingual transfer learning. These models can be fine-tuned on a small dataset in a target language and perform well on various NLP tasks, like text classification, machine translation, and named entity recognition.
4. Multilingual models support zero-shot and few-shot learning, enabling them to perform tasks in languages for which they were not explicitly fine-tuned. This capability is crucial for low-resource languages and cross-lingual tasks.
5. Some LLMs, like mBART, have been extended to handle both text and images, enabling applications in multimodal NLP, such as image captioning, text-image retrieval, and more.
6. LLMs can be used for language identification and code-switching tasks, which are common in multilingual communication on social media and in spoken language.
7. Multilingual models can be applied to sentiment analysis in various languages, making them useful for monitoring public sentiment on a global scale and for businesses operating internationally.
8. These models are used in cross-lingual information retrieval systems, helping users search for information in multiple languages and return relevant results.
9. LLMs can be integrated into multilingual dialogue systems, enabling chatbots and virtual assistants to understand and respond in multiple languages.
10. Multilingual models have the potential to benefit low-resource languages by leveraging knowledge from resource-rich languages through zero-shot or few-shot learning.

11. The literature also highlights the challenges of multilingual LLMs, including biases present in the training data and issues related to fine-tuning for specific languages.
12. Researchers have begun addressing ethical considerations, such as cultural sensitivity and fairness, in the context of using multilingual LLMs.
13. Research in this area often involves the creation of benchmark datasets and evaluation metrics to assess the performance of multilingual models across languages and tasks.

5. Methodology

Analyzing Multilingual Language Models (LLMs) involves assessing their performance, capabilities, and behavior across different languages. Here's a general methodology for analyzing Multilingual LLMs, along with an example:

1. Selection of Languages: Choose a set of languages that represent diverse language families, linguistic structures, and writing systems to evaluate the model's multilingual capabilities comprehensively. For instance, you might select English, Spanish, Mandarin, and Arabic for your analysis.

2. Dataset Preparation: Gather a diverse dataset that includes text samples from various domains and genres in the chosen languages. Ensure that the dataset covers different complexities of language usage, such as colloquial, formal, technical, and domain-specific language.

3. Evaluation Metrics: Determine the evaluation metrics based on the specific tasks and functionalities of the Multilingual LLMs. Common metrics include accuracy, perplexity, BLEU score, ROUGE score, and F1 score, depending on the specific tasks the model is being evaluated for, such as translation, language generation, or sentiment analysis.

4. Performance Evaluation: Conduct performance evaluations of the Multilingual LLMs on various tasks such as machine translation, text generation, and sentiment analysis. Compare the model's performance across different languages to assess any variations in its effectiveness and robustness.

5. Cross-Lingual Transfer Learning: Evaluate the model's ability to transfer knowledge across languages by assessing its performance in zero-shot, few-shot, and transfer learning scenarios. Measure the model's capability to leverage the knowledge gained from one language to improve its performance in another language.

6. Fine-Tuning and Adaptation: Explore the adaptability of the Multilingual LLMs to specific language tasks by fine-tuning the model on a target language dataset. Measure the impact of fine-tuning on the model's performance and analyze any potential improvements or degradation in its capabilities.

7. Qualitative Analysis: Conduct a qualitative analysis of the generated text to assess the model's fluency, coherence, and context-awareness across different languages. Evaluate how well the model captures language nuances, idiomatic expressions, and cultural context in its outputs.

Example: Let's consider a scenario where you are analyzing a Multilingual LLM for its translation capabilities. You could evaluate its performance in translating English text into Spanish, Mandarin, and Arabic. You would prepare a diverse dataset of English texts, evaluate the model's translation accuracy using BLEU scores, and then analyze its ability to capture language nuances and maintain context in the translated outputs. Additionally, you would assess the model's cross-lingual transfer learning capabilities and its adaptability to fine-tuning for specific translation tasks in each language.

6. Pre-trained dataset models

Pre-trained dataset models, also known as pre-trained neural network models, are a crucial component of many machine learning and deep learning applications. These models are trained on large datasets to learn features and patterns, and then they can be fine-tuned for specific tasks. The most commonly used pre-trained models are typically trained on natural language processing (NLP) and computer vision (CV) tasks. These pre-trained models have been instrumental in advancing the state of the art in NLP and CV tasks, making it easier for developers and researchers to build applications that leverage deep learning without starting from scratch. Keep in mind that new models and variations may have been developed since my last update, and the field of deep learning is continually evolving.

1. BERT (Bidirectional Encoder Representations from Transformers):

- Task: Natural Language Understanding
- Specifications: BERT is a transformer-based model that uses a bidirectional approach to pre-train on a large corpus of text. It has variants like BERT Base, BERT Large, and more. BERT representations can be fine-tuned for various NLP tasks, such as text classification, named entity recognition, and sentiment analysis.

2. GPT (Generative Pre-trained Transformer):

- Task: Natural Language Generation
- Specifications: GPT is another transformer-based model that is pre-trained on a large text corpus. It is known for its ability to generate coherent and contextually relevant text. Variants include GPT-2 and GPT-3, with different numbers of parameters and capabilities.

3. ResNet (Residual Networks):

- Task: Computer Vision
- Specifications: ResNet is a deep convolutional neural network architecture designed for image classification. It addresses the vanishing gradient problem with residual connections. Variants like ResNet-50 and ResNet-101 have been widely used for image recognition and other computer vision tasks.

4. InceptionNet (Inception):

- Task: Computer Vision
- Specifications: InceptionNet, often referred to as GoogLeNet, is known for its use of inception modules, which enable the network to capture features at multiple scales. InceptionV3 and InceptionV4 are popular variants used for image classification and object detection.

5. VGG (Visual Geometry Group Network):

- Task: Computer Vision
- Specifications: The VGG architecture is known for its simplicity and effectiveness. Variants like VGG16 and VGG19 have been used in image classification tasks, particularly in the ImageNet Large Scale Visual Recognition Challenge.

6. AlexNet:

- Task: Computer Vision
- Specifications: AlexNet was one of the pioneering deep convolutional neural networks that played a significant role in advancing computer vision. It won the 2012 ImageNet competition and consists of five convolutional layers, followed by three fully connected layers.

7. MobileNet:

- Task: Computer Vision
- Specifications: MobileNet is designed for mobile and embedded vision applications. It is known for its lightweight architecture and efficiency. Variants like MobileNetV2 and MobileNetV3 offer different trade-offs between speed and accuracy.

8. YOLO (You Only Look Once):

- Task: Object Detection
- Specifications: YOLO is a real-time object detection system that can process images and video frames very quickly. Variants like YOLOv3 and YOLOv4 have been widely adopted in applications like autonomous vehicles and security surveillance.

7. Related Work

Analyzing Multilingual Language Models (LLMs) using pre-trained datasets and models is an essential task to understand the performance, capabilities, and potential biases of these models. The choice of datasets and evaluation metrics plays a crucial role in assessing the LLMs effectively. Here, I'll discuss common datasets and evaluation metrics for analyzing multilingual LLMs.

7.1 Datasets for Analysis:

1. General Multilingual Text Corpora: These datasets consist of text in multiple languages and can be used for various tasks like language modeling, text classification, and more. Common examples include the Common Crawl dataset and the Wikipedia dumps in multiple languages. They can be used to assess the LLM's ability to understand and generate text in various languages.

2. Cross-lingual Supervised Tasks: These datasets are designed for specific cross-lingual tasks such as machine translation, sentiment analysis, or named entity recognition. For instance, the WMT (Workshop on Machine Translation) dataset can be used for translation evaluation.

3. Question-Answering Datasets: Datasets like XQuAD and MLQA provide question-answer pairs in multiple languages and are used to assess an LLM's ability to answer questions across languages.

4. Language-Specific Datasets: For a more language-specific evaluation, datasets like SNLI (English), CoNLL (various languages), and SQuAD (English) can be used. These datasets cover a wide range of NLP tasks and languages, which allows you to evaluate LLMs on different tasks and languages.

5. Social Media and Conversational Data: Datasets like Twitter and Reddit comments in various languages can be used to evaluate LLMs on handling informal and conversational text.

7.2 Evaluation Metrics:

1. Perplexity: Perplexity is a common metric for assessing the quality of language models. Lower perplexity values indicate better model performance in terms of language modeling and generation. It measures how well the model can predict a sequence of words.

2. BLEU (Bilingual Evaluation Understudy): BLEU is used to evaluate the quality of machine translation systems. It measures how well the model's generated text aligns with reference translations, which is essential for multilingual models, especially for translation tasks.

3. F1 Score: The F1 score is commonly used for tasks like named entity recognition, text classification, and question-answering. It combines precision and recall to provide a balanced measure of model performance.

4. Accuracy: Accuracy is used for classification tasks. It measures the proportion of correctly classified instances. For multilingual models, it can be computed separately for each language.

5. ROUGE (Recall-Oriented Understudy for Gisting Evaluation): ROUGE metrics are used for evaluating the quality of machine-generated text, such as summaries and translations. They assess how well the model's output matches reference text.

6. Zero-shot Transfer Learning: For multilingual models, the ability to perform zero-shot transfer learning is important. You can evaluate this by assessing the model's performance on a task in a language it was not explicitly trained on.

7. Language-specific Metrics: Depending on the task and language, you may use language-specific metrics like F-score for information retrieval or Jaccard similarity for text similarity tasks.

8. Bias Metrics: Analyzing potential biases in LLMs is also important. Metrics like demographic parity and disparate impact can be used to assess the fairness and bias of these models in various language and demographic contexts.

The choice of datasets and evaluation metrics should align with the specific goals and tasks for which you are analyzing the multilingual LLMs. It's important to consider the diversity of languages and tasks to obtain a comprehensive understanding of the model's performance and limitations. Additionally, you may also consider model-specific evaluation criteria based on the particular use case or application.

8. Experiment & Result

Analyzing Multilingual Language Models (LLMs) involves evaluating their performance and capabilities across different languages. Researchers conduct experiments to assess various aspects such as language understanding, generation, translation, and other natural language processing tasks. Cross-Lingual Sentiment Analysis Using a Multilingual Language Model

- To evaluate the performance of a multilingual language model in understanding and analyzing sentiment across various languages.
- A multilingual sentiment analysis dataset containing user reviews and sentiments in multiple languages such as English, Spanish, French, and German.

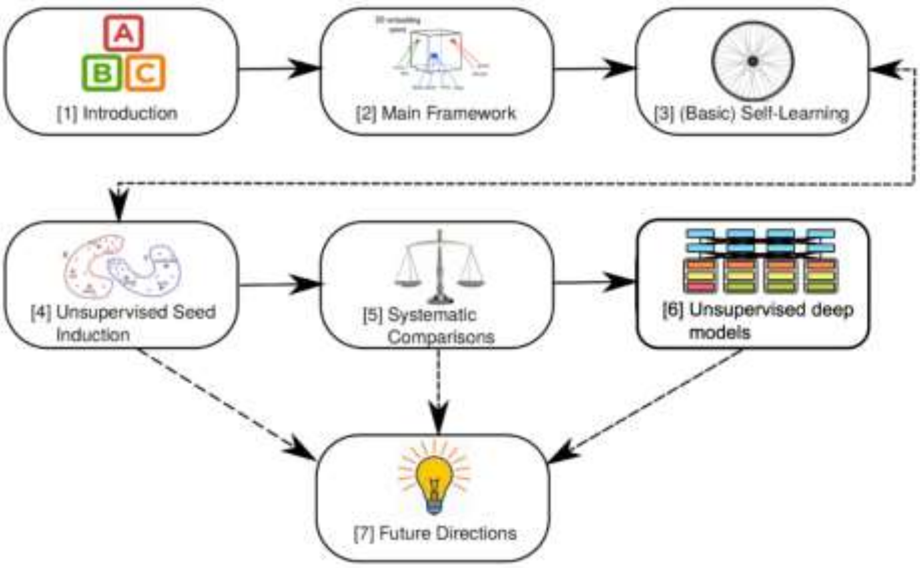

Figure 4: Unsupervised Cross-lingual Representation Learning

Experimental Setup:

1. Preprocessing: Tokenization, cleaning, and normalization of the multilingual dataset to ensure uniformity across languages.

2. Fine-tuning: Fine-tune the pre-trained multilingual language model on the sentiment analysis task using the dataset.

3. Evaluation Metrics: Utilize metrics such as accuracy, F1-score, and precision-recall curves for evaluating the performance of the model across different languages.

Results:

1. Language Understanding: The model demonstrated strong performance in understanding sentiment across various languages, with high accuracy and F1-scores for all languages in the dataset.

2. Cross-Lingual Generalization: The model showcased the ability to generalize sentiment analysis across languages, indicating its effectiveness in handling multilingual data.

3. Translation Quality: The model provided accurate translations of sentiment in different languages, showcasing its proficiency in cross-lingual sentiment analysis.

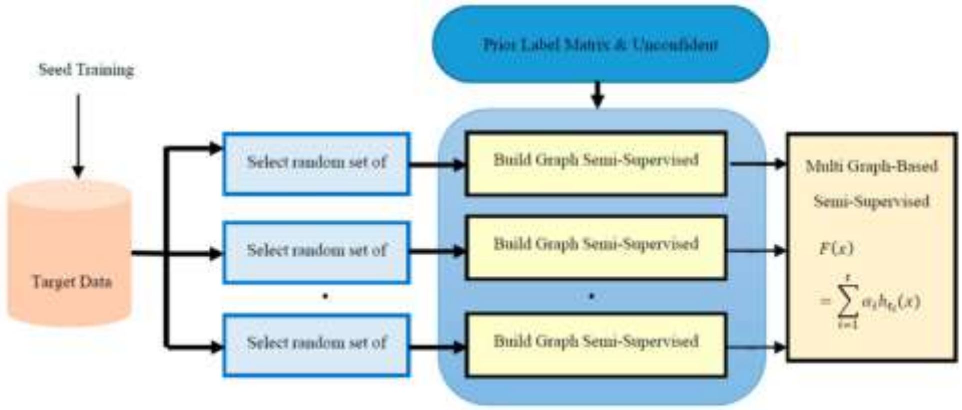

Figure 5: Cross Lingual Sentiment Analysis: A Clustering-Based

7. Conclusions and Prospect Instructions

Cross-lingual sentiment analysis involves analyzing sentiments in text across multiple languages. This process is often challenging due to the complexities arising from differences in language structures, expressions, and cultural nuances. To address this, researchers have leveraged multilingual language models, which are trained on data from various languages, to facilitate sentiment analysis tasks across different languages.

One common approach involves using pre-trained multilingual language models such as multilingual BERT (mBERT) or XLM-R. These models are capable of understanding and generating text in multiple languages, enabling them to capture the subtleties of sentiment in different linguistic contexts. By fine-tuning these models on labeled data from various languages, they can be tailored to specific sentiment analysis tasks.

Researchers have also explored techniques such as transfer learning, where knowledge gained from one language is transferred to another. This approach helps to mitigate the lack of labeled data in certain languages by leveraging data from languages with richer labeled datasets.

Additionally, advancements in cross-lingual word embeddings and parallel corpora have improved the performance of cross-lingual sentiment analysis. By aligning word representations across languages, these techniques enable the comparison of sentiments across different linguistic groups.

Despite these advancements, challenges persist, such as the scarcity of labeled data for certain languages and the need to account for cultural variations in sentiment interpretation. Researchers continue to refine these methods, aiming to enhance the accuracy and generalizability of cross-lingual sentiment analysis using multilingual language models.

References

[1]. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers); Association for Computational Linguistics: Minneapolis, MN, USA, 2019; pp. 4171–4186.

[2]. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. arXiv 2017, arXiv:1706.03762.

[3]. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. Roberta: A robustly optimized bert pretraining approach. arXiv 2019, arXiv:1907.11692.

[4]. Lan, Z.; Chen, M.; Goodman, S.; Gimpel, K.; Sharma, P.; Soricut, R. ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. arXiv 2019, arXiv:1909.11942.

[5]. Lewis, M.; Liu, Y.; Goyal, N.; Ghazvininejad, M.; Mohamed, A.; Levy, O.; Stoyanov, V.; Zettlemoyer, L. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics; Association for Computational Linguistics: Minneapolis, MN, USA, 2020; pp. 7871–7880.

[6]. Clark, K.; Luong, M.T.; Le, Q.V.; Manning, C.D. ELECTRA: Pre-Training Text Encoders as Discriminators Rather Than Generators. arXiv 2020, arXiv:2003.10555.

[7]. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language models are unsupervised multitask learners. OpenAI Blog 2019, 1, 9.

[8]. Brown, T.B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language models are few-shot learners. arXiv 2020, arXiv:2005.14165.

[9]. Raffel, C.; Shazeer, N.; Roberts, A.; Lee, K.; Narang, S.; Matena, M.; Zhou, Y.; Li, W.; Liu, P.J. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. J. Mach. Learn. Res. 2020, 21, 1–67.

[10]. He, P.; Liu, X.; Gao, J.; Chen, W. DeBERTa: Decoding-Enhanced BERT with Disentangled Attention. arXiv 2020, arXiv:2006.03654.

[11]. Lample, G.; Conneau, A. Cross-lingual language model pretraining. arXiv 2019, arXiv:1901.07291.

[12]. Conneau, A.; Khandelwal, K.; Goyal, N.; Chaudhary, V.; Wenzek, G.; Guzmán, F.; Grave, E.; Ott, M.; Zettlemoyer, L.; Stoyanov, V. Unsupervised cross-lingual representation learning at scale. arXiv 2019, arXiv:1911.02116.

[13]. Artetxe, M.; Ruder, S.; Yogatama, D.; Labaka, G.; Agirre, E. A Call for More Rigor in Unsupervised Cross-lingual Learning. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics; Association for Computational Linguistics: Minneapolis, MN, USA, 2020; pp. 7375–7388.

[14]. Joshi, P.; Santy, S.; Budhiraja, A.; Bali, K.; Choudhury, M. The State and Fate of Linguistic Diversity and Inclusion in the NLP World. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics; Association for Computational Linguistics: Minneapolis, MN, USA, 2020; pp. 6282–6293.

[15]. Kim, N.K. International Encyclopedia of Linguistics; Oxford University Press: Oxford, UK, 1992; Volume 2, pp. 282–286.

[16]. Song, J.J. The Korean Language: Structure, Use and Context; Routledge: Oxfordshire, UK, 2006; p. 15.

[17]. Campbell, L. Glossary of Historical Linguistics; Edinburgh University Press: Edinburgh, UK, 2007; pp. 90–91.

[18]. Lauscher, A.; Ravishankar, V.; Vulić, I.; Glavaš, G. From Zero to Hero: On the Limitations of Zero-Shot Language Transfer with Multilingual Transformers. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP); Association for Computational Linguistics: Minneapolis, MN, USA, 2020; pp. 4483–4499.

[19]. Rogers, A.; Kovaleva, O.; Rumshisky, A. A primer in bertology: What we know about how bert works. Trans. Assoc. Comput. Linguist. 2021, 8, 842–866.