Cross-lingual Transfer in Generative AI-Based Educational Platforms for Equitable and Personalized Learning

Nastaran Shoeibi[1]

[1] *University of Salamanca, Salamanca, Spain*

Abstract

This doctoral thesis explores the integration of Generative AI, specifically Large Language Models (LLMs) and diffusion models, in educational platforms. Emphasis is placed on cross-lingual transfer techniques to overcome language barriers and create personalized content. The study addresses the impact of Generative AI on personalized learning experiences and ethical concerns. A mixed-methods approach combines quantitative usage metrics with qualitative insights from interviews and surveys. Initial results indicate improved task performance and user engagement, but ongoing refinement is needed to address biases and ethics. The LATILL platform, a web search engine for German as Foreign Language teachers, is a case study. It leverages Generative AI to provide level-appropriate texts, translations, and image generation. The research aims to determine this technology's impact and future potential on user experience, focusing on equitable access to personalized learning across diverse geolocations.

Keywords

Generative AI, LLM, Diffusion Model, Education, Personalized Learning, Equality, language learning, Bias

1. Introduction

Personalized adaptive learning, an essential component of contemporary education, offers educational resources customized to each student's learning preferences and level of proficiency. It continuously assumes the course content and topics most relevant to a student at any given time, enhancing the learning experience [1].

Advanced AI tools like machine learning and natural language processing are used to efficiently analyze the enormous amounts of data that adaptive learning platforms generate. Also, Generative AI has improved with advances in generative adversarial networks (GANs) and deep neural networks. These advancements allow computer-generated media to resemble human-produced media, opening up opportunities in various industries, including education [2][3][4]. They can transform education by providing convenience, interactive and personalized learning experiences, and global classrooms for personalized online learning and assessment [5][6]. Large Language Models (LLMs) and diffusion models stand out as exemplars among these technologies. Diffusion models generate realistic outputs like images, videos, and audio by applying sequential transformations. In contrast, LLMs trained on large datasets excel at creating text similar to what a human would write. They are designed to handle sequential data, allowing platforms to produce original content that mimics expressions in their training data. The combination of these technologies has sped up development in various industries [7]. Even though numerous languages are widely spoken, Gehman et al [8] declare that the majority of LLMs primarily focus on monolingual skills and reflect the values of white respectability politics while giving priority to specific English dialects.

LASI: Learning Analytics Summer Institute, June 29–30, 2023, Madrid, Spain
Nastaran@usal.es (N. Shoeibi)
0000-0003-4574-6220 (N. Shoeibi)

However, this integration has challenges, such as ethical concerns, potential biases, and dissemination of inaccurate information. Addressing these issues is essential to ensuring fairness and minimizing harm. Biases can pervade models through biased training data, perpetuating societal inequalities. Vigilance in monitoring and mitigating these biases through diverse data representation and algorithmic transparency is vital. Moreover, ethical guidelines and regulations must be implemented for the responsible development and deployment of Generative AI in education and other domains [9]. When deploying these educational platforms, navigating the complicated landscape of multilingual and multicultural factors is essential. Various regions and countries have different rules and guidelines regarding data privacy, security, and algorithmic decision-making. Securing these regulations is not only a legal obligation but also a matter of ethical responsibility, ensuring the protection of students' privacy and upholding the moral principles that underpin educational endeavors [10][11].

This research endeavor provides students from diverse geolocations equal opportunities in personalized adaptive learning platforms. It explores the potential of cross-lingual transfer, multilingual LLMs, and diffusion models within Generative AI-based platforms, seeking to assess their impact on user experience. Furthermore, it examines the LATILL educational platform through a case study. Below are the research questions that will be studied:

- RQ1: How can we provide equal opportunity in personalized adaptive learning platforms for students from different geolocations?
- RQ2: Which AI Technologies empower us to do so?
- RQ3: What are the long-term impacts of using fair, personalized adaptive learning platforms?

2. Related work

Addressing data collection challenges is essential to establishing an effective educational, adaptive, personalized platform. This requires thoroughly examining works offering technical AI solutions for educational platforms. Analyzing these contributions aims to find a more seamless and personalized learning experience, ultimately benefiting students and educators alike [12].

The complexities inherent in diverse languages pose a challenge when training Large Language Models (LLMs) on multilingual datasets. However, recent advancements have led to the development of multilingual LLMs proficient in cross-lingual understanding and generation. Noteworthy examples include BLOOM [13], mBERT [14], and XLM [15], which have demonstrated strong performance in multilingual tasks.

Additionally, diffusion models have emerged as powerful tools for generating content, excelling particularly in tasks such as text-to-image synthesis and image creation. Models like DALL-E 2[2], Midjourney[3], and Stable Diffusion [4]showcase notable capabilities in this domain. Bellagente et al. introduced MULTIFUSION in [16], a system enabling the creation of intricate images by integrating various modalities and languages. MULTIFUSION effectively leverages multilingual and integrated multimodal inputs, even when trained on monomodal data in a single language. This is achieved through the fusion and alignment of pre-trained models. Furthermore, researchers in [17] propose a method known as IAP, which involves translating Stable Diffusion into Chinese. This is accomplished by refining a separate Chinese text encoder and aligning Chinese and English semantic space in CLIP.

In a related development, Cheng et al. in [18] introduce a technique to enhance the multilingual and multimodal representation model CLIP. This is achieved by substituting its text encoder with a pre-trained multilingual text encoder called XLMR.

3. Methodology

2 https://openai.com/product/dall-e-2

3 https://www.midjourney.com/

4 https://stability.ai/stable-diffusion

The methodology employed in this study is designed to effectively address the research questions and objectives, with a specific emphasis on the integration of Generative AI into educational platforms operating within multilingual contexts. A mixed-methods approach was applied to ensure a comprehensive understanding and provide meaningful insights, combining quantitative data analysis with qualitative investigation.

Firstly, quantitative data is collected by analyzing usage metrics from the educational platforms under investigation. These metrics reveal many factors, including student interactions, levels of engagement, task performance, and user preferences. The collected data undergoes different quantitative analysis techniques, including descriptive statistics and inferential analysis. These methodologies enable us to uncover patterns and trends in student behavior and platform utilization, ultimately facilitating an assessment of the effectiveness of Generative AI in enhancing personalized learning.

Additionally, qualitative insights through interviews and surveys help understand the user experience and address ethical concerns. Educators, students, and other relevant stakeholders actively participate in interviews, providing valuable perspectives on the platform's impact and any ethical dilemmas encountered. Surveys are administered to a broader sample to supplement this qualitative data, gathering additional insights into user satisfaction and concerns.

Furthermore, a central component of our methodology is using cross-lingual transfer techniques to train diffusion models and LLMs in various languages. It involves employing information from one language to improve performance in another significantly. It is typical for LLMs to pre-train on a sizable dataset in one language and fine-tune it on a smaller dataset in another. Low-resource languages no longer need as much data to be collected [19]. Similarly, diffusion models can achieve cross-lingual transfer through fine-tuning, training on multilingual datasets, or using transfer learning techniques. This lowers the data requirements while enabling the models to produce coherent and contextually relevant text in many languages. The cross-lingual transfer generally makes creating multilingual models for producing text in various languages easier [20].

4. Case study: LATILL Platform

The case study of the LATILL [5] Project demonstrates how AI technologies are employed to empower German Foreign Language (GFL) teachers in providing personalized guidance and support to students in a natural language format, which plays an important role in overcoming challenges and elevating students' educational achievements. This platform is a web search engine where teachers can select a text from the corpus and customize it with AI tools and metrics, and they can translate the selected text to several languages so the students can have it in their native language, simplify it, and generate images from scratch to visualize the selected texts for students [21][22].

- **User stories:** Through a compilation of 41 user stories from domain experts, invaluable insights were found during the platform's development journey. These reports helped identify key features and set objectives in a structured developmental strategy.
- **Platform Development:** Methodology and Technology Utilization the LATILL project's platform development part's core objective revolves around designing and implementing an educational platform tailored for GFL teachers. Functional and non-functional prototypes were designed to refine the platform's features and user interface. AI technologies, including NLP and generative AI, were integrated into the development process, enabling text translation, summarization, and image generation functionalities. This platform uses StableDiffussion for text-to-image generation, which can produce accurate and realistic images from a text prompt. Because it primarily trains on the English subset LAION2B-en of the LAION-5B dataset, it only accepts English text prompts and creates images often more oriented toward Western culture. Therefore, these prompts are translated from German to English before the photo is generated when the educator chooses a text or section of a text to

[5] Level-Adequate Texts in Language Learning

generate images. So, in some of the results, the meaning of the sentences is unrelated to the images, which is the inspiration for this paper to improve the platform. Figure 1 shows the platform's architecture.

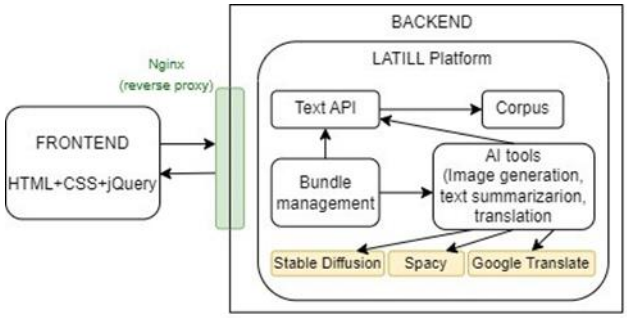

Figure 1: Architecture of the LATILL Platform

The platform's software development adhered to the SCRUM methodology [23], which champions an incremental and iterative approach. The project was compartmentalized into sprints, facilitating regular review and adaptation in response to feedback and evolving requirements. The platform's backend was architected utilizing Django, a renowned Python web framework acclaimed for its efficacy and scalability. Django formed a robust foundation for executing server-side operations, implementing business logic, and managing data.
The platform leverages state-of-the-art technology [24], such as the RTX4090 graphics card and CUDA, to expedite AI computations and graphics processing. The selection of the RTX4090 and the utilization of CUDA's parallel computing capabilities were predicated on optimizing image creation and other AI-related tasks, resulting in swift and efficient computations. Note- worthy AI management techniques, including the Stable Diffusion algorithm, Spacy, and Google Translate, facilitated image generation, simplification, and translation. These advancements underpin the platform's robust architecture and feature-rich capabilities, ensuring an optimal user experience for GFL teachers and learners.

- **User test:** After evaluating five German teachers, they noted generative AI tools cover the topic, but some key ideas are missing. They were pleased with the functionality and confirmed that this platform could be an effective resource for teaching German to students in different countries.

5. Conclusion and Future Work

In conclusion, integrating cross-lingual transfer techniques in training LLMs and diffusion models can help reduce data requirements and enhance performance across multiple languages. This advancement can facilitate societal progress, ensuring fairness, inclusivity, and equal digital access for individuals with diverse linguistic backgrounds. Moreover, implementing personalized adaptive learning within educational platforms represents a significant step towards tailoring educational experiences to individual student preferences maximizing engagement and comprehension.

The LATILL platform, equipped with advanced AI tools, is a case study showing the impact of AI on improving the user experience and assisting GFL teachers. By offering personalized materials aligned with students' interests and objectives, LATILL engages and motivates learners and addresses the specific challenges of language acquisition. While current technologies provide a strong and valuable solution, the focus in the future is to improve and integrate cross-lingual transfer in the platform and utilize Multilingual LLMs to increase user satisfaction by improving output quality and simplification results for better user understanding.

Acknowledgments

This project was undertaken with the support of the Erasmus+ Programme of the European Union: "KA2 - Cooperation partnership in school education." Level-Adequate Texts in Language Learning (LATILL) (Reference number 2021-1-AT01-KA220-SCH-000029604).

References

[1] H. Peng, S. Ma, J. M. Spector, Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment, Smart Learning Environments 6 (2019).
[2] T. Karras, S. Laine, T. Aila, A style-based generator architecture for generative adversarial networks, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4401–4410.
[3] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE transactions on pattern analysis and machine intelligence 40 (2017) 834–848.
[4] F. J. García-Peñalvo, Developing robust state-of-the-art reports: Systematic literature reviews, Education in the Knowledge Society 23 (2022) e28600. doi:10.14201/eks. 28600.
[5] C. Yang, C. Chen, W. He, R. Cui, Z. Li, Robot learning system based on adaptive neural control and dynamic movement primitives, IEEE transactions on neural networks and learning systems 30 (2018) 777–787.
[6] P. Gao, J. Li, S. Liu, An introduction to key technology in artificial intelligence and big data driven e-learning and e-education, Mobile Networks and Applications 26 (2021) 2123–2126.
[7] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, et al., A survey of large language models, arXiv preprint arXiv:2303.18223 (2023).
[8] S. Gehman, S. Gururangan, M. Sap, Y. Choi, N. A. Smith, RealToxicityPrompts: Evalu- ating neural toxic degeneration in language models, in: Findings of the Association for Computational Linguistics: EMNLP 2020, Association for Computational Linguistics, 2020.
[9] S. Harrer, Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine, eBioMedicine 90 (2023) 104512.
[10] J. Borenstein, A. Howard, Emerging challenges in ai and the need for ai ethics education, AI and Ethics 1 (2021) 61–65.
[11] Nguyen, H. N. Ngo, Y. Hong, B. Dang, B.-P. T. Nguyen, Ethical principles for artificial intelligence in education, Education and Information Technologies 28 (2023) 4221–4241.
[12] F. Ouyang, P. Jiao, Artificial intelligence in education: The three paradigms, Computers and Education: Artificial Intelligence 2 (2021) 100020.
[13] Workshop, :, T. L. Scao, A. Fan, C. Akiki, E. Pavlick, S. Ilić, D. Hesslow, etc., Bloom: A 176b-parameter open-access multilingual language model, 2023. arXiv:2211.05100.
[14] H. Gonen, S. Ravfogel, Y. Elazar, Y. Goldberg, It's not greek to mbert: inducing word-level translations from multilingual bert, arXiv preprint arXiv:2010.08275 (2020).
[15] Z. Chi, S. Huang, L. Dong, S. Ma, B. Zheng, S. Singhal, P. Bajaj, X. Song, X.-L. Mao,
[16] H. Huang, et al., Xlm-e: Cross-lingual language model pre-training via electra (2021).
[17] M. Bellagente, M. Brack, H. Teufel, F. Friedrich, B. Deiseroth, C. Eichenberg, A. Dai,
[18] R. Baldock, S. Nanda, K. Oostermeijer, et al., Multifusion: Fusing pre-trained models for multi-lingual, multi-modal image generation, arXiv preprint arXiv:2305.15296 (2023).
[19] J. Hu, X. Han, X. Yi, Y. Chen, W. Li, Z. Liu, M. Sun, Efficient cross-lingual transfer for chinese stable diffusion with images as pivots, 2023. arXiv:2305.11540.
[20] Z. Chen, G. Liu, B.-W. Zhang, F. Ye, Q. Yang, L. Wu, Altclip: Altering the language encoder in clip for extended language capabilities, 2022. arXiv:2211.06679.

[21] M. Artetxe, H. Schwenk, Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond, Transactions of the Association for Computational Linguistics 7 (2019) 597–610.
[22] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolution image synthesis with latent diffusion models, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10684–10695.
[23] Z. Soproni, Common european framework of reference for languages: Learning, teaching, assessment, companion volume with new descriptors, Modern Nyelvoktatás 26 (2020) 168–170.
[24] M. Kienberger, A. García-Holgado, K. Schramm, A. Raveling, D. Meurers, B. Labinska, T. Koropatnitska, R. Therón, Enhancing Adaptive Teaching of Reading Skills Using Digital Technologies: The LATILL Project, 2023, pp. 1092–1098. doi:10.1007/978-981-99-0942-1_ 115.
[25] L. Rising, N. S. Janoff, The scrum software development process for small teams, IEEE software 17 (2000) 26–32.
[26] R. Zhang, D. Zou, Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning, Computer Assisted Language Learning 35 (2022) 696–742.