JT-IBSI (Jurnal Teknik Ibnu Sina), Vol 9, No. 02, Desember 2024
E-ISSN: 2541-2647 DOI: https://doi.org/10.36352/jt-ibsi.v9i02.1061

UTILIZATION OF AI IN NATURAL LANGUAGE PROCESSING (NLP) A LITERATURE REVIEW

[1]Aprizal Y, [2]Romiko Afriantoni, [3]Willy Rizki Perdana, [4]Ismail, [5]Taslimahudin, [6]Hanafi, [7]Revi illya Badri, [8]Atman Lucky Fernandes
[1,4,5,6,7,8]Program Studi Teknik Informatika, Fakultas Teknik –Universitas Ibnu Sina, Batam
[2,3]Program Studi Sistem Informasi, Fakultas Teknik –Universitas Ibnu Sina, Batam
e-mail: afrizal@uis.ac.id

Abstract

The utilization of Artificial Intelligence (AI) in Natural Language Processing (NLP) has advanced rapidly over the past five years. Transformer-based models such as BERT, GPT-4, T5, and LLaMA-2 have significantly improved NLP's ability to understand, analyze, and generate human language with increasing accuracy. These technologies have been applied across various sectors, including industry, healthcare, education, and public services, through virtual assistants, chatbots, automated translation, and social media sentiment analysis. However, several challenges remain in NLP development, such as the limited representation of low-resource languages, biases in models, and high computational costs. Additionally, ethical and privacy concerns in AI-powered NLP applications are critical issues. Therefore, innovation is needed in developing more energy-efficient models, strategies to mitigate bias, and stricter data protection policies. By adopting a more inclusive, transparent, and sustainable approach, AI-driven NLP can provide broader benefits to society across various domains..

Keywords: *Artificial Intelligence, Natural Language Processing, NLP, Transformer, AI Ethics, AI Inclusivity*

Accepted : July 2024
Approved : August 2024
Published : December 2024

Introduction

The development of artificial intelligence (AI) technology has brought revolutionary changes across various fields, one of which is Natural Language Processing (NLP). NLP is a branch of AI that focuses on the understanding, analysis, and generation of human language by computers (Jurafsky & Martin, 2021). With advancements in Machine Learning and Deep Learning, NLP has undergone significant progress, enabling various sophisticated applications such as machine translation, sentiment analysis, chatbots, and virtual assistants (Young et al., 2018).

The advancement of NLP technology is closely linked to the development of Transformer-based models, such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer), which have opened new possibilities in understanding the structure and meaning of text more effectively (Vaswani et al., 2017). These models can handle long-term dependencies in text, comprehend broader contexts, and improve accuracy in various natural language processing tasks (Devlin et al., 2019).

Today, NLP is widely used in various sectors, including industry, healthcare, education, and public services (Hirschberg & Manning, 2015). Some increasingly prevalent NLP applications include virtual assistants (Google Assistant, Siri, Alexa), customer service chatbots, and automated text analysis systems (Radford et al., 2019). In the healthcare sector, NLP is used for analyzing electronic medical records, assisting doctors in disease diagnosis, and providing treatment recommendations (Wu et al., 2020). Additionally, NLP plays a crucial role in speech recognition, text classification, and social media analysis to identify public opinion (Cambria et al., 2017).

Despite the rapid advancements in AI-based NLP, several challenges remain. One major issue is bias in language models, where AI models can reproduce or even amplify biases present in their training data (Bender et al., 2021). Furthermore, NLP still faces difficulties in understanding complex contexts, requires massive amounts of training data, and has limitations in handling low-resource languages (Blodgett et al., 2020).

Another challenge in NLP utilization relates to ethics and privacy. The use of NLP models for analyzing private texts, conversations, or sensitive data poses risks to information security and data misuse (Floridi & Cowls, 2019). Therefore, further research is needed to ensure that NLP is developed with considerations of transparency, fairness, and social responsibility (Weidinger et al., 2021).

Based on this background, this study aims to examine the utilization of AI in NLP through a literature review approach. By analyzing various existing studies, this research will discuss the latest trends in NLP, ongoing challenges, and its potential applications across different sectors. The findings of this study are expected to provide deeper insights for academics, industry practitioners, and technology developers in understanding and optimizing AI-based NLP for various future needs.

Research Method

This study employs a literature review approach to analyze the utilization of AI in Natural Language Processing (NLP). The research process involves identifying, selecting, and reviewing relevant academic sources, including journal articles, conference papers, and technical reports from the past five years. The primary focus is on advancements in transformer-based models such as BERT, GPT-4, T5, and LLaMA-2, along with their applications, challenges, and future directions.

The data collection process follows a systematic literature review method, utilizing reputable databases such as ACM Digital Library, and Google Scholar. Selected studies are categorized based on key themes, including model development, ethical concerns, computational efficiency, and practical implementations across various sectors.

A qualitative analysis is conducted to synthesize findings, compare existing methodologies, and highlight emerging trends in NLP research. By critically evaluating prior works, this study aims to provide insights into how AI-driven NLP technologies evolve, the limitations they face, and potential innovations for broader, ethical, and inclusive applications.

Results and Discussion

Results

1. Advancements in AI for Natural Language Processing (NLP) (2019–2024)

Over the past five years, Artificial Intelligence (AI) has brought significant transformations to the field of Natural Language Processing (NLP). Advances in transformer models such as BERT (Bidirectional Encoder Representations from Transformers), T5 (Text-to-Text Transfer Transformer), GPT-4, and the latest LLaMA-2 have demonstrated substantial improvements in natural language processing (Touvron et al., 2023).

Compared to traditional methods such as n-gram and rule-based approaches, transformer-based deep learning models have a superior ability to understand context through the self-attention mechanism. As a result, NLP has become more accurate in grasping linguistic nuances, generating more natural text, and enhancing various applications such as machine translation, chatbots, text summarization, and sentiment analysis (Brown et al., 2020).

Recent Models and Their Advantages:

- BERT (2019): Enables bidirectional word context understanding, improving accuracy in text classification and information extraction (Devlin et al., 2019).
- GPT-3 & GPT-4 (2020–2023): Capable of generating more natural text and understanding context in AI-driven dialogues (OpenAI, 2023).
- LLaMA-2 (2023): An open-source model that is more efficient and energy-saving, developed by Meta AI (Touvron et al., 2023).

2. Challenges in AI Implementation for NLP

Although AI in NLP has advanced significantly, several challenges still need to be addressed, including:

a. Limitations in Language Representation

Most modern NLP models are dominated by languages with abundant data, such as English and Mandarin, while low-resource languages still struggle to achieve accurate representation (Joshi et al., 2020).

For instance, Indonesian language processing faces challenges due to the lack of high-

quality datasets. Efforts to improve this have been made through projects like IndoBERT and other transformer-based models specifically developed for local languages (Wilie et al., 2020).

b. Bias in NLP Models

NLP models often reflect biases present in the datasets used for training. A study by Bender et al. (2021) found that AI models frequently exhibit gender, racial, and social biases, which can negatively impact applications such as sentiment analysis, automated recruitment, and customer service chatbots.

To mitigate these biases, various approaches have been developed, such as DebiasBERT, which aims to reduce bias in text classification (Liang et al., 2022).

c. High Computational Costs

State-of-the-art AI models, such as GPT-4 and PaLM-2, require enormous computational resources. Patterson et al. (2022) found that training advanced NLP models can consume thousands of GPUs and terabytes of RAM, as well as generate high carbon emissions.

Solutions to These Challenges:

- Quantization: Reducing model precision to save computational power.
- Pruning: Removing less important parameters from the NLP model.
- Federated Learning: Training models in a decentralized manner to reduce central computational demands.

a. Model Interpretability (Black Box Problem)

Most NLP models still function as black boxes, meaning it is difficult to understand how they make decisions. This poses challenges in fields such as law, healthcare, and finance, where interpretability is crucial (Rudin, 2019).

Approaches like SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) have been introduced to enhance transparency in NLP models.

3. Applications of AI in NLP Across Various Sectors

AI-powered NLP has been applied in multiple fields, including:

a. Healthcare

- Medical Record Analysis, NLP aids in extracting information from electronic health records (EHR) for automated diagnosis (Esteva et al., 2021).
- Healthcare Chatbots, Applications like Ada Health and Babylon Health use NLP to respond to users' medical inquiries.

b. Education

- AI-Based Tutoring Systems, NLP is used in developing automated tutors that can answer student questions and provide personalized learning (Sahu et al., 2022).
- Plagiarism Detection, NLP algorithms help detect plagiarism in academic work by comparing text against existing sources.

c. Business and Finance

- Stock Market Sentiment Analysis, NLP analyzes news and social media to predict market trends (Chollet, 2023).
- Legal Document Automation, AI processes legal documents automatically, reducing lawyers' workload.

4. Implications and Future Directions of AI-Based NLP

Based on research findings, several key trends are expected to shape the future of NLP over the next five years:

- More Efficient and Eco-Friendly NLP Models
  The development of low-power AI models with parameter optimization and reduced carbon footprint.
- Improved Interpretability and Fairness
  New methods to reduce bias in AI-powered NLP and enhance model transparency (Bommasani et al., 2022).
- Expansion to Low-Resource Languages
  Few-shot learning and multilingual models will help improve NLP support for underrepresented languages (Touvron et al., 2023).
- Integration of NLP with Multimodal AI
  Future models will combine text, images, and speech, as seen in OpenAI's GPT-4 and Google's Gemini models.

Discussion

Artificial Intelligence (AI) has brought rapid advancements in Natural Language Processing (NLP) over the past five years. Transformer-based models such as BERT, GPT-4, T5, and LLaMA-2 have significantly improved NLP's ability to understand, analyze, and generate human language with greater accuracy. This technology has been applied across various sectors, including industry, healthcare, education, and public services, in the form of virtual assistants, chatbots, automatic translation, and sentiment analysis on social media.

However, there are still major challenges in NLP development, including the limitations of low-resource language representation, biases in models, and high computational costs.

Additionally, ethical and privacy concerns remain key issues in AI-driven NLP applications. Therefore, innovations are needed in developing more energy-efficient models, bias mitigation strategies, and stricter data protection policies. With a more inclusive, transparent, and sustainable approach, AI-based NLP can provide broader benefits to society across various fields.

Conclusion

AI in NLP has rapidly advanced in the past five years, driven by transformer models like BERT, GPT-4, and T5, enhancing language understanding and generation across various sectors. However, challenges remain, including biases, high computational costs, and limited support for resource-scarce languages. Efforts to develop energy-efficient models, reduce bias, and promote linguistic inclusivity are ongoing. Additionally, ethical and privacy concerns highlight the need for stricter regulations to ensure responsible AI implementation.

Reference

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? . Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT), 610–623.

Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of “bias” in NLP. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 5454–5476.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., & Bernstein, M. S. (2022). On the opportunities and risks of foundation models. Journal of Artificial Intelligence Research, 65, 23–45.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems (NeurIPS), 33, 1877–1901.

Cambria, E., White, B., Huang, G., & Hussain, A. (2017). SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1).

Chollet, F. (2023). Deep Learning with Python. Manning Publications.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 4171–4186.

Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., & Topol, E. (2021). Deep learning-enabled medical computer vision. Nature Medicine, 27(1), 14–22.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1).

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. Science, 349(6245), 261–266.

Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity in the NLP world. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 6282–6293.

Jurafsky, D., & Martin, J. H. (2021). Speech and Language Processing (3rd ed.). Prentice Hall.

Liang, P., Chen, Y., & Zou, W. (2022). DebiasBERT: An empirical study on mitigating bias in transformer-based text classification. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP).

OpenAI. (2023). GPT-4 technical report. arXiv preprint arXiv:2303.08774.

Patterson, D., Gonzalez, J., Le, Q. V., Liang, C., & Dean, J. (2022). The carbon impact of large-scale AI training. Communications of the ACM, 65(11), 66–75.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2019). Improving language understanding by generative pre-training. OpenAI Technical Report.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215.

Sahu, A., Gupta, R., & Srivastava, N. (2022). AI-based tutoring systems: A survey on applications and future directions. International Journal of Artificial Intelligence in Education, 32(3), 455–478.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., & Scialom, T. (2023). LLaMA: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems (NeurIPS), 30, 5998–6008.

Weidinger, L., Mellor, J., Rauh, M., Griffin, D., & Whittlestone, J. (2021). Ethical and societal implications of AI-based NLP. AI Ethics Journal, 5(4), 315–333.

Wilie, B., Vincentio, K. R., Winata, G. I., Cahyawijaya, S., Li, H., Lim, Z., & Bahar, R. (2020). IndoBERT: A pre-trained model for Indonesian NLP. arXiv preprint arXiv:2009.05387.

Wu, S., Roberts, K., Datta, S., Du, J., Ji, Z., Si, Y., & Xu, H. (2020). Deep learning in clinical natural language processing: A methodical review. Journal of the American Medical Informatics Association (JAMIA), 27(3), 457–470.

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning-based natural language processing. IEEE Computational Intelligence Magazine, 13(3), 55–75.