SJER Scientific Journal of Engineering Research

PTTI PENELITI TEKNOLOGI TEKNIK INDONESIA

Date of publication January 3, 2025, date of current version January 3, 2025.
Digital Object Identifier ***10.59247/sjer.v1i1.6***

Review

A Bibliometric Analysis of Natural Language Processing and Classification: Trends, Impact, and Future Directions

Setiawan Ardi Wijaya[1,*], Rahmad Gunawan[2,3], Rangga Alif Faresta[4], Asno Azzawagama Firdaus[5], Gabriel Diemesor[6], Furizal[7]

[1] Department of Information System, Universitas Muhammadiyah Riau, Pekanbaru, 28294, Indonesia; setiawanardiwijaya@umri.ac.id
[2] Faculty of Data Science and Computing, Universiti Malaysia Kelantan, City Campus, Pengkalan Chepa, 16100 Kota Bharu, Kelantan, Malaysia; goengoen78@umri.ac.id
[3] Department of Informatics Engineering, Faculty of Computer Sciences, Universitas Muhammadiyah Riau, Pekanbaru, Indonesia; goengoen78@umri.ac.id
[4] Department of Digital Learning, Monash University, Clayton Victoria, Australia; rfar0018@student.monash.edu
[5] Department of Computer Science, Universitas Qamarul Huda Badaruddin Bagu, Central Lombok, 83371, Indonesia; asnofirdaus@uniqhba.ac.id
[6] Department of Chemical Engineering, University of Benin, Benin, 320001, Nigeria; diemesorgabriel@gmail.com
[7] Department of Informatics Engineering, Universitas Islam Riau, Pekanbaru, 28284, Indonesia; furizal.id@gmail.com
* Correspondence

This research was supported by Department of Information System, Universitas Muhammadiyah Riau. We also want to thank Lembaga Penelitian & Pengabdian kepada Masyarakat (LPPM) Universitas Muhammadiyah Riau. Our gratitude is also to Universitas Muhammadiyah Riau for other support.

Abstract: This study presents a bibliometric analysis of Natural Language Processing (NLP) and classification re-search, examining trends, impacts, and future directions. NLP, a key field in artificial intelligence, focuses on enabling computers to process and understand human language through tasks such as text classification, sentiment analysis, and speech recognition. Classification plays a crucial role in organizing textual data, facilitating applications like spam detection and content recommendation. The research employs bibliometric analysis to evaluate publication trends, citation networks, and emerging themes from 1992 to 2025. Using data retrieved from Scopus, descriptive statistical analysis and bibliometric mapping with VOSviewer reveal key contributors, influential publications, and subject area distributions. Findings indicate a significant rise in NLP research, with deep learning models, particularly transformers, driving advancements in the field. The study highlights dominant research areas, including computer science, engineering, and medicine, and identifies leading countries in NLP research, such as the United States, China, and India. Additionally, ethical concerns, including bias and fairness in NLP applications, are discussed as critical challenges for future research. The insights derived from this analysis provide valuable guidance for researchers and policymakers in shaping the next phase of NLP development.

Keywords: Trend analysis; Natural Language Processing; Classification; Bibliometric Analysis; VOSviewer

1. Introduction

The study of artificial intelligence's Natural Language Processing (NLP) area is concerned with how computers and human language interact. It entails the creation of models and algorithms that allow computers to comprehend, interpret, and produce meaningful and practical human language. NLP is a broad field that includes tasks like speech recognition, text classification, sentiment analysis, and language translation, all of which are intended to close the gap between human and machine understanding. NLP technologies can improve user experiences, automate tasks, and extract insights from massive amounts of text data. These technologies can be applied to a wide range of applications, including search engines and virtual assistants [1-4].

NLP and classification work hand in hand, with classification being a key method used in NLP to arrange and

analyze textual data. Classification tasks in NLP entail grouping text according to predetermined labels, like identifying a document's topic, assessing a review's sentiment (positive, negative, or neutral), or identifying named entities like people, places, and organizations. Applications such as spam detection, language translation, and content recommendation are made possible by classification models, which are frequently trained on sizable, labeled datasets. These models allow NLP systems to make precise predictions and judgments based on the textual input. This relationship is essential because classification aids in the organization of unstructured text data, improving its usability and accessibility for a range of language-based technologies [5-8].

The field of NLP and classification has made significant progress in recent times due to the improved accuracy and scalability of language processing tasks. One of the most notable trends in NLP and classification is the extensive use of deep learning models such as transformers. These developments have had a significant impact, with NLP and classification technologies now being essential to a wide range of applications, from sentiment analysis and predictive text to chat-bots and automated customer support. Future developments in the field are probably going to concentrate on improving model interpretability and fairness, dealing with biases, and extending NLP capabilities to handle increasingly complex and context-rich language tasks. These developments will have a lasting impact on how NLP and classification are integrated into a wider range of industries, posing both new opportunities and difficulties for the creation of language-based technologies [9-12].

In the context of NLP and classification, bibliometric analysis is essential because it offers a methodical and quantitative way to evaluate the state of the re-search, identify important trends, significant contributors, and developing fields of study. Bibliometric analysis maps the field's evolution by examining publications, citations, and collaborations. This helps to identify the most influential research and the paths taken by scholarly communication. Understanding how NLP and classification have changed over time, what subjects are becoming more popular, and how research is dispersed among various institutions and geographical areas are all made possible by this kind of analysis. Furthermore, by highlighting gaps, encouraging cooperation, and influencing funding and policy decisions, bibliometric insights can direct future research and accelerate the development of natural language processing and classification technologies. This study aims so answer these research question:

RQ1: What are current Natural Language Processing and Classification research and publication trends?

RQ2: Which is the subject area that involves research and publication in Natural Language Processing and Classification?

RQ3: What are the 10 most prominent source titles that contribute in terms of total publication and citations in Natural Language Processing and Classification?

RQ4: What are the most highly cited documents in Natural Language Processing and Classification research?

RQ5: Which keywords and countries are most contributing to Natural Language Processing and Classification research?

2. Literature Review

The development of Natural Language Processing (NLP) has evolved from rule-based systems and statistical methods to advanced machine learning and deep learning approaches[13]. With the advent of machine learning, supervised models became more prevalent, enabling tasks like sentiment analysis and document categorization to achieve higher levels of accuracy [14]. The introduction of deep learning and neural networks marked a turning point in NLP, as demonstrated by the success of recurrent neural networks (RNNs) and later, transformer models such as BERT [15-18] and GPT [19-22]. By using cutting-edge contextualization techniques, these models have dramatically improved machines' capacity to comprehend and produce human language, revolutionizing the field.

Beyond technological developments, NLP and classification technologies pre-sent ethical challenges that have been highlighted in recent literature. Concerns about bias, fairness, and model interoperability have received more attention in research, particularly as NLP systems are incorporated into crucial processes like hiring and legal decision-making [23-25]. A few scholars highlight the necessity of accountability and transparency in AI models, encouraging the industry to adopt more responsible development procedures [26-28]. These discussions are crucial as they address the societal implications of NLP systems, ensuring that advancements in technology do not come at the cost of fairness and equity. As the field progresses, the integration of ethical considerations into the research and deployment of NLP systems will become increasingly essential.

3. Methodology

The methodology used in this research is presented in Figure 1. The flowchart illustrates the process of data collection and retrieval for a study on "Natural Language Processing" and "Classification." The process begins with the authors determining the topic of interest and selecting the timeframe for data retrieval. A search query using `TITLE-ABS-KEY("Natural Language Processing" AND "Classification")` is performed, resulting in the retrieval of 255 documents from 1992 onwards. The data is then retrieved in CSV format for further processing. Following this, the retrieved data is divided into two formats for analysis. One format, Scopus_exported_refine_values.csv, is processed using Microsoft Excel to conduct Descriptive statistics and analyze publication trends. The other format, Scopus.csv, is analyzed using VOSviewer for bibliometric

mapping. This structured approach ensures that the retrieved data is comprehensively analyzed using different tools and methodologies to extract meaningful insights.

4. Results and Discussion

4.1. Trends in Publications

Figure 2 shows a significant rise in publications and citations from 1992 to 2025, with major peaks in 2019, 2021, and especially 2022. This indicates growing research interest and increasing impact, highlighting the topic's growing relevance in recent years.

4.2. Publication by Countries

Figure 3 shows global citation relationships, with the United States, China, and India as central hubs, indicating their significant influence and collaboration in research. Countries like South Korea, Germany, and the UK are also active, while Saudi Arabia and Kazakhstan are more isolated. The map highlights strong global collaborations, with key countries leading in citation impact.

4.3. Publications by Subject Area

Table 1 shows highlight the distribution of publications across various subject areas related to natural language processing research. Computer Science leads significantly with 67.45% of the total publications, followed by Engineering (36.47%) and Medicine (26.27%), indicating that NLP research is heavily concentrated in technical fields. Other notable areas include Mathematics (16.86%) and Decision Sciences (12.94%), showing interdisciplinary applications of NLP. Fields like Social Sciences (9.80%) and Neuroscience (2.35%) have comparatively fewer contributions, suggesting a more limited integration of NLP in these areas. Overall, the data emphasizes NLP's primary focus on computer science and engineering, with growing relevance in medicine and other disciplines.

Table 1. Publications by Subject Area.

Subject Area	TP	Percentage
Computer Science	172	67.45%
Engineering	93	36.47%
Medicine	67	26.27%
Mathematics	43	16.86%
Decision Sciences	33	12.94%
Social Sciences	25	9.80%
Physics and Astronomy	18	7.06%
Materials Science	14	5.49%
Energy	11	4.31%
Health Professions	8	3.14%
Business, Management and Accounting	7	2.75%
Neuroscience	6	2.35%
Biochemistry, Genetics and Molecular Biology	5	1.96%

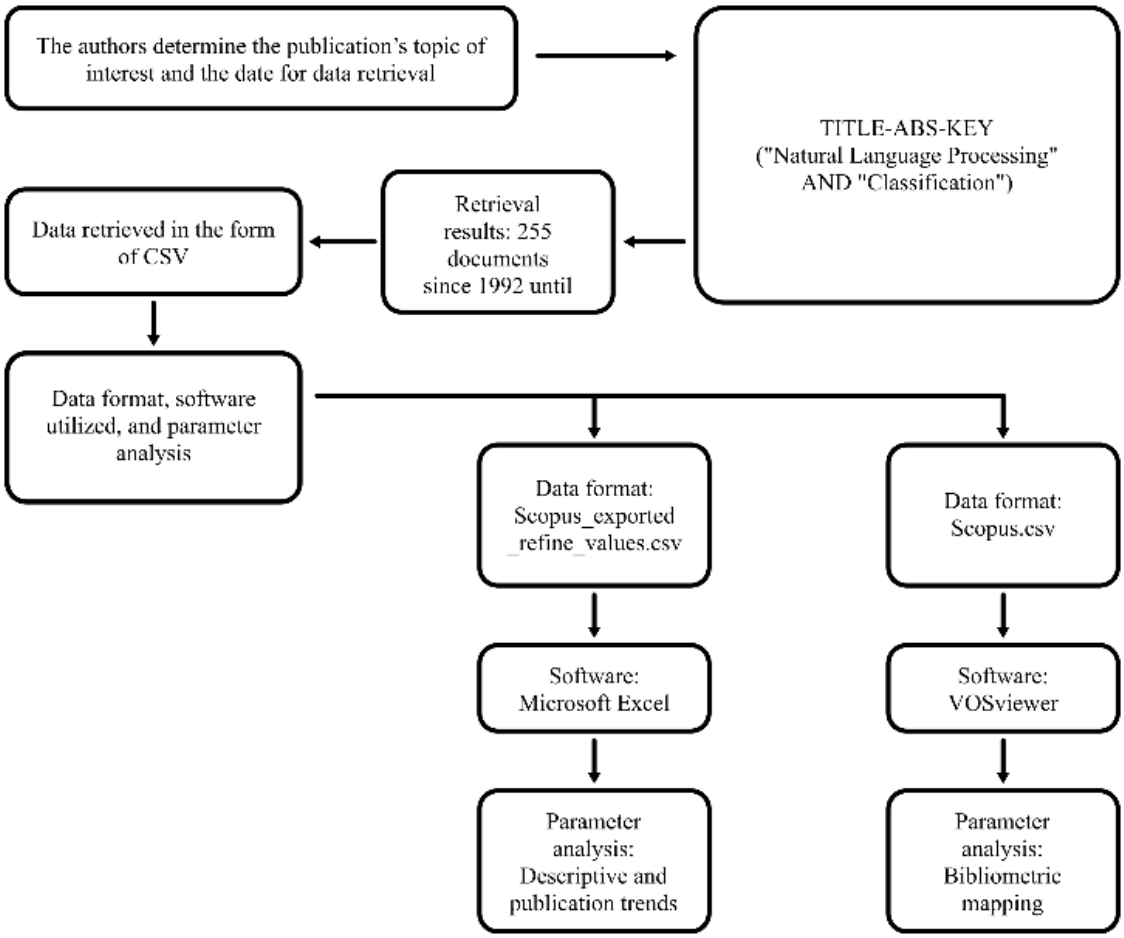

Figure 1. The Flowchart of Data Collection and Data Retrieval.

Table 2. Most active source titles.

Source Title	TP	NCA	NCP	TC	C/P	C/CP	h	g	m
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	7	32	6	24	3.43	4.00	4	4	0.333
Lecture Notes in Networks and Systems	5	21	3	5	1.00	1.67	2	2	0.667
Journal of Medical Internet Research	5	22	3	4	0.80	1.33	1	1	0.250

Notes: TP=total number of publications; NCP=number of cited publications; TC=total citations; C/P=average citations per publication; C/CP=average citations per cited publication; h=h-index; and g=g-index.

Table 3. Top 3 highly cited articles.

No.	Author(s)	Title	Source Title	TC	C/Y
1	Tanguy L.; Tulechki N.; Urieli A.; Hermann E.; Raynal C. (2016)	Natural language processing for aviation safety reports: From classification to interactive analysis	Computers in Industry	132	14.67
2	Weng W.-H.; Wagholikar K.B.; McCray A.T.; Szolovits P.; Chueh H.C. (2017)	Medical subdomain classification of clinical notes using a machine learning-based natural language processing approach	BMC Medical Informatics and Decision Making	103	12.88
3	Garg R.; Oh E.; Naidech A.; Kording K.; Prabhakaran S. (2019)	Automating Ischemic Stroke Subtype Classification Using Machine Learning and Natural Language Processing	Journal of Stroke and Cerebrovascular Diseases	96	16.00

4.4. Publications by Sources Titles and Highly Cited Documents

Table 2 compares three different sources based on their publication and citation metrics. Lecture Notes in Computer Science, with seven total publications (TP), has the highest number of citations per publication (C/P = 3.43) and a strong citation per cited publication (C/CP = 4.00). It also has the highest h-index (h = 4), indicating consistent influence. Lecture Notes in Networks and Systems has five publications and a lower citation rate (C/P = 1.00), though its citation per cited publication is slightly higher (C/CP = 1.67). Lastly, Journal of Medical Internet Research shows a modest impact with five publications and the lowest citation rate per publication (C/P = 0.80). Overall, *Lecture Notes in Computer Science* demonstrates greater academic impact in terms of citations and publication influence.

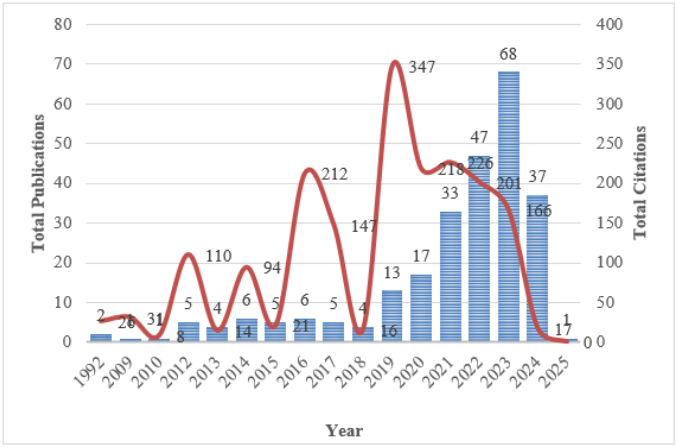

Figure 2. Total Publications and Citations by Year.

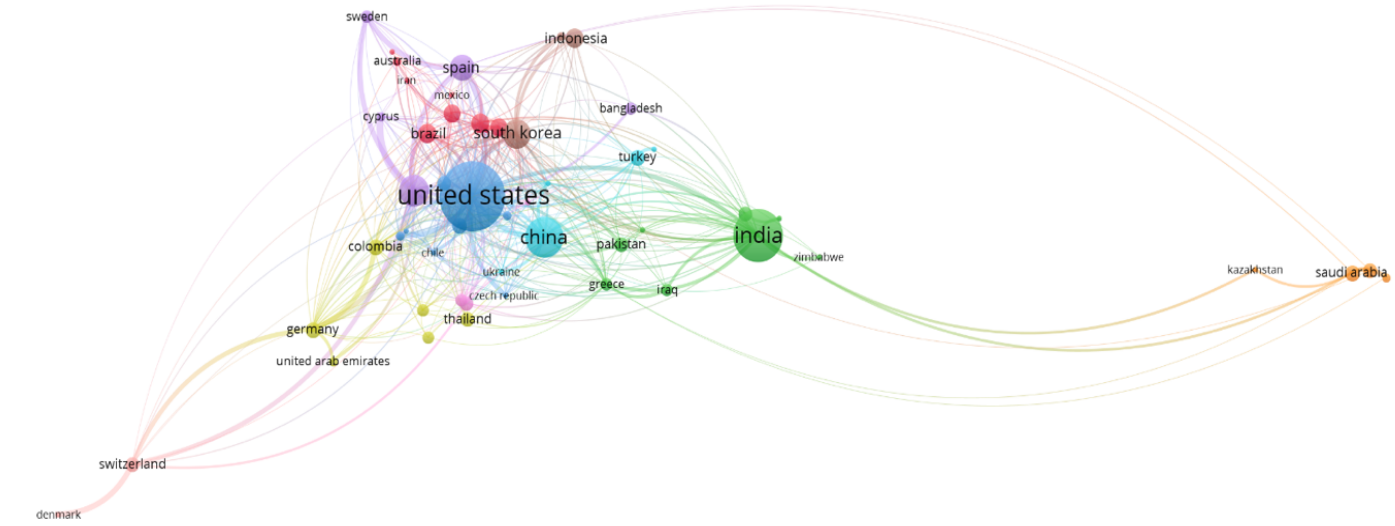

Figure 3. Network visualization map of the citations based on countries.

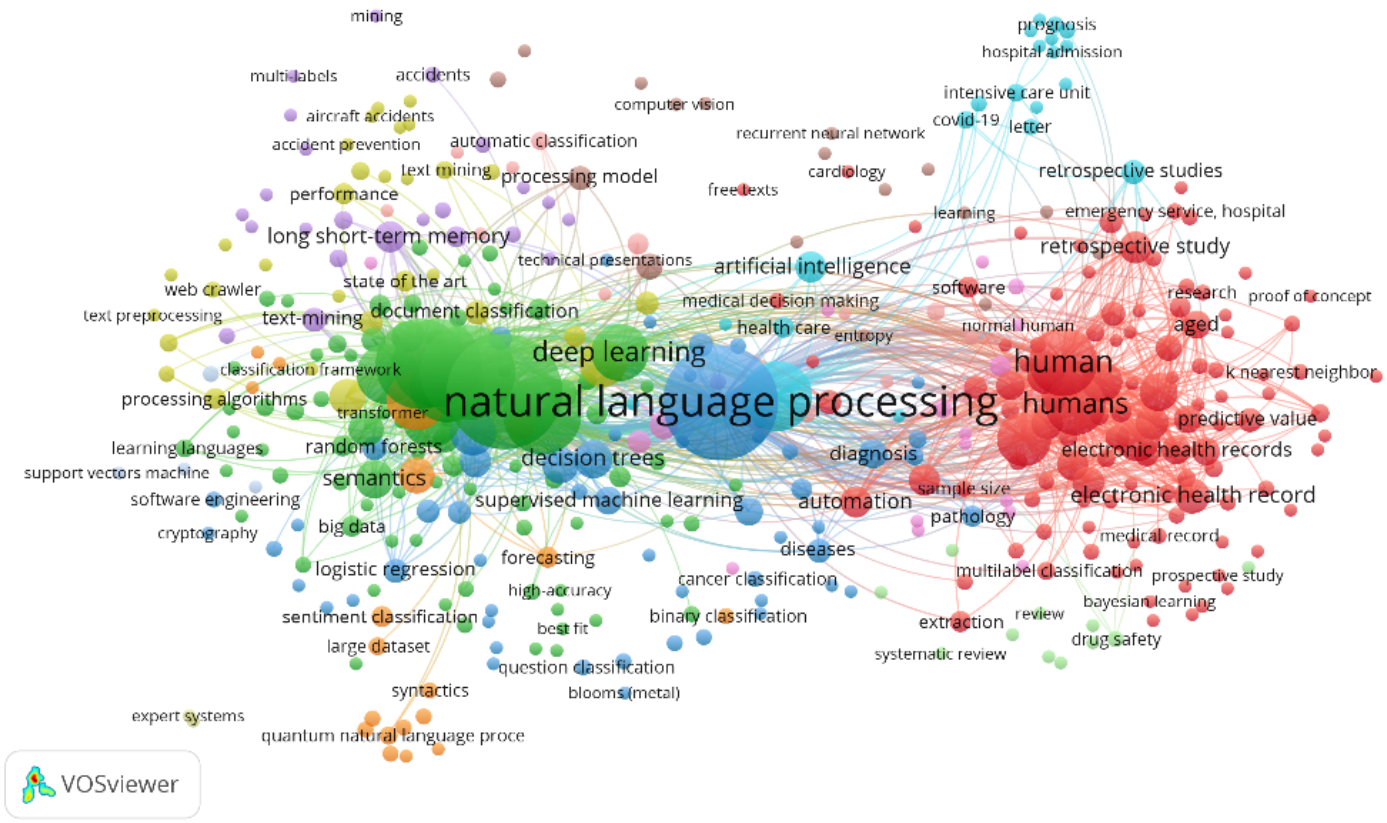

Figure 4. Network visualization map of the co-occurrence by keywords

Table 3 presents three studies on the application of natural language processing (NLP) across different domains. The study by Tanguy et al. (2016) on aviation safety reports has the highest total citations (TC = 132) with an average of 14.67 citations per year. Weng et al. (2017) applied NLP for classifying medical subdomains in clinical notes, with 103 total citations and an average of 12.88 citations per year. Garg et al. (2019) focused on ischemic stroke subtype classification using NLP and machine learning, with the fewest total citations (96) but the highest citation rate per year (16.00), reflecting its recent impact. These studies highlight the diverse applications and strong academic influence of NLP in specialized fields.

4.5. Most Used Keyword

Figure 4 shows the network visualization map created by VOSviewer shows the co-occurrence of keywords related to "Natural Language Processing" (NLP). The map is color-coded, with clusters representing different research themes. The central theme is NLP, linked to areas like deep learning, machine learning, and artificial intelligence (in green and blue), while other clusters (in red) focus on applications in healthcare, such as electronic health records and human-related studies. The interconnected nodes reveal relationships between terms, indicating major trends and collaborations in the research field.

5. Conclusion

The study of Natural Language Processing (NLP) focuses on enabling computers to understand and process human language through various tasks such as speech recognition, text classification, and sentiment analysis. NLP relies heavily on classification models to organize and analyze textual data, which allows for applications like spam detection, content recommendation, and language translation. With advancements in deep learning, particularly the use of transformer models, NLP has seen significant improvements in accuracy and scalability. These developments have led to widespread adoption of NLP technologies in diverse fields, including predictive text, sentiment analysis, and automated customer support. However, ethical concerns such as bias and fairness in NLP models have gained attention as these systems become integrated into critical decision-making processes, highlighting the importance of transparency and accountability.

The study uses bibliometric analysis to explore trends in NLP and classification research, identifying key areas of growth, influential publications, and major contributors. It reveals an increasing number of publications and citations over the years, with countries like the United States, China, and India leading the re-search. NLP research is predominantly concentrated in technical fields such as computer science and engineering, with growing applications in medicine and decision sciences. The analysis also highlights top research sources and highly cited articles, which demonstrate the diverse application of NLP across fields like aviation safety, healthcare, and stroke classification. Major keywords in the field center around deep learning, artificial intelligence, and healthcare applications, reflecting the field's evolving trends and collaborative nature.

6. Conflicts of Interest

The author (s) declare no conflict of interest.

7. References

[1] S. Park, X. Wang, C. C. Menassa, V. R. Kamat, and J. Y. Chai, "Natural language instructions for intuitive human interaction with robotic assistants in field construction work," *Autom Constr*, vol. 161, May 2024, doi: 10.1016/j.autcon.2024.105345.

[2] F. Wang, A. C. K. Cheung, and C. S. Chai, "Language learning development in human-AI interaction: A thematic review of the research landscape," *System*, vol. 125, Oct. 2024, doi: 10.1016/j.system.2024.103424.

[3] J. Barbosa *et al.*, "Evaluating the noise tolerance of Cloud NLP services across Amazon, Microsoft, and Google," *Comput Ind*, vol. 164, Jan. 2025, doi: 10.1016/j.compind.2024.104211.

[4] M. H. Kazemi and A. Alvanchi, "Application of NLP-based models in automated detection of risky contract statements written in complex script system," *Expert Syst Appl*, vol. 259, Jan. 2025, doi: 10.1016/j.eswa.2024.125296.

[5] A. Hur, N. Janjua, and M. Ahmed, "Unifying context with labeled property graph: A pipeline-based system for comprehensive text representation in NLP," *Expert Syst Appl*, vol. 239, Apr. 2024, doi: 10.1016/j.eswa.2023.122269.

[6] Y. Wu and J. Wan, "A survey of text classification based on pre-trained language model," *Neurocomputing*, vol. 616, Feb. 2025, doi: 10.1016/j.neucom.2024.128921.

[7] Q. Xi and P. Jiang, "Design of news sentiment classification and recommendation system based on multi-model fusion and text similarity," *International Journal of Cognitive Computing in Engineering*, vol. 6, pp. 44–54, Dec. 2025, doi: 10.1016/j.ijcce.2024.11.003.

[8] Y. Wang, C. Gong, X. Ji, and Q. Yuan, "Text classification for evaluating digital technology adoption maturity based on BERT: An evidence of Industrial AI from China," *Technol Forecast Soc Change*, vol. 211, Feb. 2025, doi: 10.1016/j.techfore.2024.123903.

[9] R. S. Abdul Kareem, T. Tilford, and S. Stoyanov, "Fine-grained food image classification and recipe extraction using a customized deep neural network and NLP," *Comput Biol Med*, vol. 175, Jun. 2024, doi: 10.1016/j.compbiomed.2024.108528.

[10] G. Huang, Y. Li, S. Jameel, Y. Long, and G. Papanastasiou, "From explainable to interpretable deep learning for natural language processing in healthcare: How far from reality?," Dec. 01, 2024, *Elsevier B.V.* doi: 10.1016/j.csbj.2024.05.004.

[11] M. Arslan and C. Cruz, "Leveraging NLP approaches to define and implement text relevance hierarchy framework for business news classification," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 317–326. doi: 10.1016/j.procs.2023.10.016.

[12] S. A. Beecher Martins, N. Garrido, and P. Sebastião, "Port request classification automation through NLP," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 1927–1934. doi: 10.1016/j.procs.2024.06.376.

[13] K. D. Amin *et al.*, "Development and Validation of a Natural Language Processing Model to Identify Low-Risk Pulmonary Embolism in Real Time to Facilitate Safe Outpatient Management," *Ann Emerg Med*, vol. 84, no. 2, pp. 118–127, Aug. 2024, doi: 10.1016/j.annemergmed.2024.01.036.

[14] W. Zhu, J. Qiu, Z. Yu, and W. Luo, "A survey on personalized document-level sentiment analysis," *Neurocomputing*, vol. 609, Dec. 2024, doi: 10.1016/j.neucom.2024.128449.

[15] P. Dhiman, A. Kaur, D. Gupta, S. Juneja, A. Nauman, and G. Muhammad, "GBERT: A hybrid deep learning model based on GPT-BERT for fake news detection," *Heliyon*, vol. 10, no. 16, Aug. 2024, doi: 10.1016/j.heliyon.2024.e35865.

[16] H. Bekamiri, D. S. Hain, and R. Jurowetzki, "PatentSBERTa: A deep NLP based hybrid model for patent distance and classification using augmented SBERT," *Technol Forecast Soc Change*, vol. 206, Sep. 2024, doi: 10.1016/j.techfore.2024.123536.

[17] S. Xu, C. Zhang, and D. Hong, "BERT-based NLP techniques for classification and severity modeling in basic warranty data study," *Insur Math Econ*, vol. 107, pp. 57–67, 2022, Accessed: Jan. 03, 2025. [Online]. Available: https://doi.org/10.1016/j.insmatheco.2022.07.013

[18] M. S. Mithun *et al.*, "BERT NLP MODEL FOR MULTICLASS CLASSIFICATION OF RADIOLOGY REPORTS," *Physica Medica*, vol. 104, p. S52, Dec. 2022, doi: 10.1016/s1120-1797(22)02236-0.

[19] S. Haroon, C. A. Hafsath, and A. S. Jereesh, "Generative Pre-trained Transformer (GPT) based model with relative attention for de novo drug design," *Comput Biol Chem*, vol. 106, Oct. 2023, doi: 10.1016/j.compbiolchem.2023.107911.

[20] F. Sufi, "Advanced Computational Methods for News Classification: A Study in Neural Networks and CNN integrated with GPT," *Journal of Economy and Technology*, Sep. 2024, doi: 10.1016/j.ject.2024.09.001.

[21] D. Zhang *et al.*, "Utilizing GPT-4 for CT Image Analysis in Cerebral Hemorrhage: Innovating Applications of Natural Language Processing in Radiology (Preprint)," *J Med Internet Res*, Sep. 2024, doi: 10.2196/58741.

[22] T. J. Gracie, Y. Pershad, C. Bejan, A. G. Bick, B. Ferrell, and A. Kishtagari, "Combined Natural Language Processing and Gpt-4 Pathology Report Interpretation Efficiently Identify a Myelodysplastic Syndrome Cohort for Large Scale Clinical Research Applications," *Blood*, vol. 144, no. Supplement 1, pp. 3607–3607, Nov. 2024, doi: 10.1182/blood-2024-201083.

[23] M. Javaid, A. Haleem, and R. P. Singh, "A study on ChatGPT for Industry 4.0: Background, potentials, challenges, and eventualities," *Journal of Economy and Technology*, vol. 1, pp. 127–143, Nov. 2023, doi: 10.1016/j.ject.2023.08.001.

[24] P. Sawant and K. Sonawane, "NLP-based smart decision making for business and academics," *Natural Language Processing Journal*, p. 100090, Jul. 2024, doi: 10.1016/j.nlp.2024.100090.

[25] P. Sawant and K. Sonawane, "NLP-based smart decision making for business and academics," *Natural Language Processing Journal*, vol. 8, no. 100090, 2024, Accessed: Jan. 03, 2025. [Online]. Available: https://doi.org/10.1016/j.nlp.2024.100090

[26] B. Memarian and T. Doleck, "Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review," Jan. 01, 2023, *Elsevier B.V.* doi: 10.1016/j.caeai.2023.100152.

[27] J. Bagenal *et al.*, "Generative AI: ensuring transparency and emphasising human intelligence and accountability," Nov. 30, 2024, *Elsevier B.V.* doi: 10.1016/S0140-6736(24)02615-1.

[28] B. Memarian and T. Doleck, "Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review," Jan. 01, 2023, *Elsevier B.V.* doi: 10.1016/j.caeai.2023.100152.