

NATURAL LANGUAGE PROCESSING

DALAM SISTEM INFORMASI DAN KECERDASAN BUATAN



Tar Muhammad Raja Gunung, S.Kom., M.Kom.

Siti Sahara Lubis, S.Kom., M.Kom.

Yerezyq Bagus, S.Kom., M.Kom.

Seta Samsiana ST., MT.

Salsabila Miftah Rezkia, S.T., M.Sc.

Nurul Mega Saraswati, M.Kom.

Manutur Pandapotan Siregar, S.T., M.Kom.

Natural Language Processing dalam Sistem Informasi dan Kecerdasan Buatan

Tar Muhammad Raja Gunung, S.Kom., M.Kom.

Siti Sahara Lubis, S.Kom., M.Kom.

Yerezqy Bagus, S.Kom., M.Kom.

Seta Samsiana ST., MT.

Salsabila Miftah Rezkia, S.T., M.Sc.

Nurul Mega Saraswati, M.Kom.

Manutur Pandapotan Siregar, S.T., M.Kom.

PT BUKULOKA LITERASI BANGSA

Anggota IKAPI: No. 645/DKI/2024



***Natural Language Processing* dalam Sistem Informasi dan Kecerdasan Buatan**

Penulis : Tar Muhammad Raja Gunung, S.Kom., M.Kom., Siti Sahara Lubis, S.Kom., M.Kom., Yerezqy Bagus, S.Kom., M.Kom., Seta Samsiana ST., MT., Salsabila Miftah Rezkia, S.T., M.Sc., Nurul Mega Saraswati, M.Kom., dan Manutur Pandapotan Siregar, S.T., M.Kom.

ISBN : 978-634-324-038-9 (PDF)

Penyunting Naskah : Fahrul Rozi

Tata Letak : Fahrul Rozi

Desain Sampul : Novikean Keysah Sanisri

Penerbit

Penerbit PT Bukuloka Literasi Bangsa

Distributor: PT Yapindo

Kompleks Business Park Kebon Jeruk Blok I No. 21, Jl. Meruya Ilir Raya No. 88, Desa/Kelurahan Meruya Utara, Kecamatan Kembangan, Kota Adm. Jakarta Barat, Provinsi DKI Jakarta, Kode Pos: 11620

Email: penerbit.blb@gmail.com

Whatsapp: 0878-3483-2315

Website: bukuloka.com

© Hak cipta dilindungi oleh undang-undang

Berlaku selama 50 (lima puluh) tahun sejak ciptaan tersebut pertama kali dilakukan pengumuman.

Dilarang mengutip atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari penerbit. Ketentuan Pidana Sanksi Pelanggaran Pasal 2 UU Nomor 19 Tahun 2002 Tentang Hak Cipta.

Barang siapa dengan sengaja dan tanpa hak melakukan prebuttal sebagaimana dimaksud dalam Pasal 2 ayat (1) atau Pasal 49 ayat (1) dan ayat (2) dipidana dengan pidana penjara masing-masing paling singkat 1 (satu) bulan dan/atau denda paling sedikit Rp1.000.000,00 (satu juta rupiah), atau pidana penjara paling lama 7 (Tujuh) tahun dan/atau denda paling banyak Rp5.000.000.000,00 (lima miliar rupiah).

Barang siapa dengan sengaja menyerahkan, menyiarkan, memamerkan, mengedarkan atau menjual kepada umum suatu ciptaan atau barang hasil pelanggaran Hak Cipta atau Hak Terkait sebagaimana dimaksud pada ayat (1) dipidana dengan pidana penjara paling lama 5 (lima) tahun dan/atau denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya, sehingga buku referensi berjudul ***Natural Language Processing dalam Sistem Informasi dan Kecerdasan*** dapat disusun dan diterbitkan. Buku ini hadir untuk memberikan pemahaman mengenai pengolahan bahasa alami sebagai salah satu bidang yang berperan penting dalam perkembangan sistem informasi dan teknologi kecerdasan masa kini.

Pembahasan dalam buku ini disusun secara ringkas dan mudah dipahami, mencakup konsep-konsep utama, proses kerja, serta penerapan pengolahan bahasa alami dalam berbagai kebutuhan digital. Materi yang disajikan diharapkan dapat membantu pembaca mengenal bagaimana komputer memahami, memproses, dan menghasilkan bahasa yang digunakan manusia.

Buku ini diterbitkan untuk masyarakat umum dan dipasarkan secara luas agar dapat diakses oleh siapa saja yang memiliki minat terhadap perkembangan teknologi informasi dan kecerdasan. Semoga buku ini memberikan manfaat, menambah wawasan, serta mendorong pembaca untuk lebih memahami pemanfaatan teknologi bahasa dalam kehidupan sehari-hari.

Jakarta, Juni 2026

Tim Penyusun

DAFTAR ISI

KATA PENGANTAR.....	iii
DAFTAR ISI.....	iv
Bab 1: Konsep Dasar <i>Natural Language Processing</i>.....	1
1.1 Pengertian NLP	1
1.2 Sejarah dan Perkembangan NLP	5
1.3 Ruang Lingkup NLP	9
1.4 Tantangan dalam Pemrosesan Bahasa Alami.....	14
1.5 Peran NLP dalam Sistem Cerdas.....	17
Bab 2: Representasi Teks dan Data Bahasa	21
2.1 Karakteristik Data Teks.....	21
2.2 Tokenisasi dan Normalisasi Teks.....	23
2.3 <i>Stopword</i> dan <i>Stemming</i>	26
2.4 Representasi Berbasis Vektor (BoW, TF-IDF)	28
2.5 Embedding Kata (Word Embedding).....	31
Bab 3: Linguistik Komputasi.....	35
3.1 Struktur Bahasa (Fonologi, Morfologi, Sintaksis, Semantik)	35
3.2 Analisis Morfologi	39
3.3 Mengurai Sintaksis	43
3.4 Analisis Semantik.....	47
3.5 Peran Linguistik dalam NLP	50
Bab 4: Pemrosesan Teks Dasar.....	55
4.1 Preprocessing Teks.....	55
4.2 Pembersihan Data Teks	64
4.3 Transformasi Teks.....	70

4.4 Ekstraksi Fitur Teks.....	77
4.5 Pipeline Pemrosesan Teks	85
Bab 5: Model Bahasa (Language Model).....	94
5.1 Konsep Model Bahasa.....	94
5.2 N-gram Model	97
5.3 Probabilitas dalam NLP.....	100
5.4 Evaluasi Model Bahasa	103
5.5 Keterbatasan Model Tradisional	106
Bab 6: Pembelajaran Mesin untuk NLP.....	110
6.1 Konsep Machine learning dalam NLP	110
6.2 Klasifikasi Teks	115
6.3 Clustering Teks.....	119
6.4 Feature engineering	123
6.5 Evaluasi Model ML untuk Teks.....	127
Bab 7: Deep Learning dalam <i>Natural Language Processing</i>..	133
7.1 Neural Network untuk Teks	133
7.2 Recurrent neural network (RNN)	136
7.3 LSTM dan GRU	138
7.4 Transformer	141
7.5 Fine-Tuning Model NLP	144
Bab 8: Tugas-Tugas Utama NLP.....	147
8.1 Penandaan Bagian Ucapan	147
8.2 Pengenalan Entitas Bernama (NER)	150
8.3 Analisis Sentimen.....	153
8.4 Klasifikasi Teks	157
8.5 Terjemahan Mesin	159
Bab 9: Sistem Pencarian Informasi dan Teks	163
9.1 Konsep Information Retrieval	163

9.2 Indexing dan Ranking	166
9.3 Query processing	169
9.4 Evaluasi Sistem Pencarian.....	172
9.5 Aplikasi Pencarian Teks.....	174
Bab 10: Evaluasi dan Validasi Model NLP	178
10.1 Metode Evaluasi NLP	178
10.2 Precision, Recall, dan F1-Score	181
10.3 Cross-Validation.....	185
10.4 Evaluasi Berbasis Dataset	189
10.5 Interpretasi Hasil Model.....	192
Bab 11: Penerapan NLP dalam Sistem.....	197
11.1 Integrasi NLP dalam Aplikasi	197
11.2 NLP Berbasis Api dan Framework	200
11.3 Chatbot dan Virtual assistant.....	203
11.4 NLP dalam Sistem Informasi	205
11.5 Arsitektur Sistem NLP	208
Bab 12: Pengembangan dan Integrasi NLP Berkelanjutan...	211
12.1 Strategi Pengembangan Sistem NLP.....	211
12.2 Integrasi NLP dengan Sistem Lain.....	214
12.3 Manajemen Data dan Model NLP.....	217
12.4 Etika dan Bias dalam NLP	220
12.5 Optimalisasi dan Pemeliharaan Sistem NLP	223
Daftar Pustaka.....	227
Profil Penulis	239

Daftar Pustaka

- Abubakar, H. D., & Umar, M. (2022). Sentiment Classification: Review of Text Vectorization Methods: Bag of Words, Tf-Idf, Word2vec and Doc2vec. *SLU Journal of Science and Technology*. <https://doi.org/10.56471/slujst.v4i.266>
- Agarkar, Dr. P. M., Kumar, A., & Sahay, M. (2025). Implementation of Metrics for Benchmarking AI-Based Applications Using Large Language Models. *International Journal of Advanced Research in Science, Communication and Technology*. <https://doi.org/10.48175/ijarsct-30481>
- Ai, Q., Bi, K., Guo, J., & Croft, W. B. (2018). Learning a Deep Listwise Context Model for *Ranking* Refinement. Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. <https://doi.org/10.1145/3209978.3209985>
- Banerjee, P., & Kashyap, S. (2024). Unlocking Transfer Learning's Potential in *Natural Language Processing*: An Extensive Investigation and Evaluation. 2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET). <https://doi.org/10.1109/ACROSET62108.2024.10743260>
- Chai, C. P. (2022). *Comparison of text preprocessing methods*. Natural Language Engineering.

Chai, K. M. A. (2022). *Preprocessing Methods for Natural Language Processing: Tokenization, Normalization, and Noise Removal Techniques*. Information Processing & Management, 59(5), 103111.

Chai, Y., Jin, L., Feng, S., & Xin, Z. (2024). Evolution and advancements in deep learning models for *Natural Language Processing*. Applied and Computational Engineering. <https://doi.org/10.54254/2755-2721/77/20240674>

Deekshith, A. (2024). ADVANCES IN NATURAL LANGUAGE PROCESSING: A SURVEY OF TECHNIQUES. International Journal of Innovations in Engineering Research and Technology. <https://doi.org/10.26662/ijiert.v8i3.pp74-83>

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. NAACL-HLT 2019 Proceedings, 4171–4186.
DOI: <https://doi.org/10.48550/arXiv.1810.04805>

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: Human Language Technologies, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>

DOI: <https://doi.org/10.1016/j.ipm.2022.103111>

DOI: <https://doi.org/10.1017/S1351324922000213>

DOI: <https://doi.org/10.70767/jmec.v1i2.257>

Gardner, M., Grus, J., Neumann, M., Tafjord, O., Dasigi, P., Liu, N. F., Peters, M. E., Schmitz, M., & Zettlemoyer, L. (2018). *AllenNLP: A Deep Semantic Natural Language Processing Platform*. arXiv.Org. <https://doi.org/10.18653/v1/W18-2501>

Gusenbauer, M., & Haddaway, N. (2019). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Wiley. <https://doi.org/https://doi.org/10.1002/jrsm.1378>

Ho, C. F. (2025). The Declining Role of Lemmatization and Stemming in Modern *NLP* Pipelines. 2025 9th International Conference on Man-Machine Systems (ICoMMS). <https://doi.org/10.1109/ICoMMS66553.2025.11200245>

Hu, Y., & Ghadimi, P. (2024). In-Depth Analysis of *Recall* Initiators of Medical Devices with a *Machine learning-Natural Language Processing* Tool. International Conference on Informatics in Control, Automation and Robotics. <https://doi.org/10.5220/0012900600003822>

Husada, M. S., & J.p, S. Y. (2026). Analisis Emosi dalam Lirik Lagu menggunakan *Natural Language Processing*. Jurnal Indonesia : Manajemen Informatika Dan Komunikasi. <https://doi.org/10.63447/jimik.v7i1.1697>

Jabbar, A., Iqbal, S., Tamimy, M. I., Rehman, A., Bahaj, S. A. O., & Saba, T. (2023). *An Analytical Analysis of Text Stemming Methodologies in Information Retrieval and Natural*

- Language Processing Systems*. IEEE Access.
DOI: <https://doi.org/10.1109/ACCESS.2023.3332710>
- Jabbar, A., Khan, S., & Mahmood, T. (2023). *Stemming and Lemmatization Techniques in Natural Language Processing: A Comparative Study*. *Applied Sciences*, 13(9), 5482.
DOI: <https://doi.org/10.3390/app13095482>
- K, C., S, R., S, S., & M, S. (2026). Optimizing Text Classification Accuracy using *Machine learning* and Deep Learning Algorithms. 2026 6th International Conference on Image Processing and Capsule Networks (ICIPCN).
<https://doi.org/10.1109/ICIPCN67432.2026.11438448>
- Kamath, B. P., Member, I. U. D. M. G., Acharya, D., Ayush, S., Rao, S., Rai, A. R., & Shetty. (2025). Comprehensive Analysis of *Word Embedding* Models and Design of Effective Feature Vector for Classification of Amazon Product Reviews. IEEE Access. <https://doi.org/10.1109/ACCESS.2025.3536631>
- Khan, Md. R. (2024). Advancements in Deep Learning Architectures: A Comprehensive Review of Current Trends. *Journal of Artificial Intelligence General Science (JAIGS)* ISSN:3006-4023. <https://doi.org/10.60087/jaigs.v1i1.29>
- Kumar, A., Singh, P., & Verma, S. (2023). *Noise Reduction and Text Cleaning Strategies for Social Media Data Analytics*. *Journal of Big Data*, 10(1), 88.
DOI: <https://doi.org/10.1186/s40537-023-00788-2>
- Kumar, P. (2024). Large language models (LLMs): survey, technical frameworks, and future challenges. Springer

Science+Business

Media.

<https://doi.org/https://doi.org/10.1007/s10462-024-10888-y>

Lin, J., Ma, X., Lin, S.-C., Yang, J.-H., Pradeep, R., & Nogueira, R. (2021). *Pyserini: A Python Toolkit for Reproducible Information Retrieval Research with Sparse and Dense Representations*. None.

<https://doi.org/https://doi.org/10.1145/3404835.3463238>

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2022). *Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing*. Association for Computing Machinery.

<https://doi.org/https://doi.org/10.1145/3560815>

Malamon, J. (2024). *DNA N-gram analysis framework (DNAnamer): A generalized N-gram frequency analysis framework for the supervised classification of DNA sequences*. bioRxiv.

<https://doi.org/10.1016/j.heliyon.2024.e36914>

Mienye, I. D., Swart, T. G., & Obaido, G. (2024). *Recurrent neural networks: A Comprehensive Review of Architectures, Variants, and Applications*. Inf.

<https://doi.org/10.3390/info15090517>

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. International Conference on Learning Representations (ICLR).

DOI: <https://doi.org/10.48550/arXiv.1301.3781>

- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv:1301.3781*. <https://doi.org/10.48550/arXiv.1301.3781>
- Mujtaba, G., Shuib, N., Idris, N., Hoo, W. L., Raj, R. G., Khowaja, K., Shaikh, K., & Nweke, H. F. (2019). Clinical text classification research trends: Systematic literature review and open issues. *Expert Systems with Applications*. <https://doi.org/10.1016/J.ESWA.2018.09.034>
- Nguyen, M.-A., Sai, K., & Nguyen, M.-A. (2026). Tracing the Evolution of *Word Embedding* Techniques in *Natural Language Processing*. <https://www.semanticscholar.org/paper/5608146cfcf1c1d758564f9183a9943112e3b517>
- Ni, J., Pu, J., Yang, Z., Zhou, K., Wang, H., Xiao, X., Wang, D., Li, X., Luo, J., & Hu, C. (2025). From Large to Super-Tiny: End-to-End Optimization for Cost-Efficient LLMs. *arXiv.Org*. <https://doi.org/10.48550/arXiv.2504.13471>
- Ohno, Y., Aomori, T., Nishiyama, T., Kato, R., Fujiki, R., Ishikawa, H., Kiyomiya, K., Isawa, M., Mochizuki, M., Aramaki, E., & Ohtani, H. (2024). Performance Improvement of a *Natural Language Processing* Tool for Extracting Patient Narratives Related to Medical States From Japanese Pharmaceutical Care Records by Increasing the Amount of Training Data: *Natural Language Processing Analysis and Validation Study*. *JMIR Medical Informatics*. <https://doi.org/10.2196/68863>

- Opitz, J., Wein, S., & Schneider, N. (2024). *Natural Language Processing RELIES on Linguistics*. Computational Linguistics. https://doi.org/10.1162/coli_a_00560
- Pakray, P., Gelbukh, A., & Bandyopadhyay, S. (2025). *Natural Language Processing applications for low-resource languages*. *Natural Language Processing*. <https://doi.org/10.1017/NLP.2024.33>
- Paleyes, A., Urma, R.-G., & Lawrence, N. D. (2022). Challenges in Deploying *Machine learning*: A Survey of Case Studies. Association for Computing Machinery. <https://doi.org/https://doi.org/10.1145/3533378>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). *Scikit-learn: Machine learning in Python*. *Journal of Machine learning Research*, 12, 2825–2830. DOI: <https://doi.org/10.48550/arXiv.1201.0490>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). *Scikit-learn: Machine learning in Python*. *Journal of Machine learning Research*, 12(85), 2825–2830.
- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. *Proceedings of the 2014*

Conference on Empirical Methods in Natural Language Processing, 1532–1543. <https://doi.org/10.3115/v1/D14-1162>

Pradhana, R. K., Seerapu, D. P., Routhu, G., Manda, S. K., & Jami, G. R. (2025). Locally Deployed *NLP* System for Secure Document Summarization and Context-Aware Question Answering Using LLMs and Vector Embeddings. *International Journal for Sciences and Technology*. <https://doi.org/10.71097/ijسات.v16.i2.4275>

Sampaio, T., Oliveira, P., & Matos, P. (2025). 3D-Chatbot: A Intelligent Conversational Agent Integrating RAG, *NLP*, And VR For Commercial Stores. 2025 International Conference on Engineering and Emerging Technologies (ICEET). <https://doi.org/10.1109/ICEET67911.2025.11424080>

Satwika, C. G. S., & Pramod, J. (2024). Exploring Word-Level Representations in Modern *Natural Language Processing*. *International Journal for Research in Applied Science and Engineering Technology*. <https://doi.org/10.22214/ijraset.2024.64791>

Sen, A. (2025). Efficiency and Performance Trade-Offs: A Comparative Analysis of Statistical N-Gram Models and Resource-Optimized Small Language Models (SLMs) for Edge Computing Applications. *International Journal of Innovative Science and Research Technology*. <https://doi.org/10.38124/ijisrt/25nov395>

- Sheikhalishahi, S., Miotto, R., Dudley, J. T., Lavelli, A., Rinaldi, F., & Osmani, V. (2019). *Natural Language Processing of Clinical Notes on Chronic Diseases: Systematic Review*. JMIR Publications.
<https://doi.org/https://doi.org/10.2196/12239>
- Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). *Text Data Augmentation for Deep Learning*. Springer Science+Business Media.
<https://doi.org/https://doi.org/10.1186/s40537-021-00492-0>
- Tonello, N., Macdonald, C., & Ounis, I. (2018). *Efficient Query processing for Scalable Web Search*. Foundations and Trends in Information Retrieval.
<https://doi.org/10.1561/15000000057>
- Tunstall, L., von Werra, L., & Wolf, T. (2022). *Natural Language Processing with Transformers: Building language applications with Hugging Face*. O'Reilly Media.
- Tyagi, A., Jain, V., & Kumar, V. (2024). Benchmark Text Preprocessing Techniques in *Natural Language Processing*. 2024 4th International Conference on Innovative Sustainable Computational Technologies (CISCT).
<https://doi.org/10.1109/CISCT62494.2024.11134188>
- Tyagi, A., Jain, V., & Kumar, V. (2024). *Benchmark Text Preprocessing Techniques in Natural Language Processing*. 2024 4th International Conference on Innovative Sustainable Computational Technologies (CISCT). DOI: <https://doi.org/10.1109/CISCT62494.2024.11134188>

- Tyagi, V., Sharma, R., & Gupta, P. (2024). *Comparative Analysis of NLP Preprocessing Techniques Using NLTK, SpaCy, and Transformer Tokenizers*. *Expert Systems with Applications*, 245, 122931.
DOI: <https://doi.org/10.1016/j.eswa.2024.122931>
- Vanin, D. (2025). The Application of Monolingual and Multilingual BERT-Based Models for Text Automation Tasks. *NaUKMA Research Papers Computer Science*.
<https://doi.org/10.18523/2617-3808.2025.8.76-83>
- Vanita. (2024). An Extant of *Natural Language Processing*. *International Journal For Multidisciplinary Research*.
<https://doi.org/10.36948/ijfmr.2024.v06i02.18326>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. *Advances in Neural Information Processing Systems* (NeurIPS), 5998–6008.
DOI: <https://doi.org/10.48550/arXiv.1706.03762>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- Vijayaragavan, Dr. K., Selvan, T., Sudarsan, R., & Sreeraam, R. K. (2024). Integrating *NLP Chatbot* into Career Guidance Web Application. *International Journal of Advanced Research in Science, Communication and Technology*.
<https://doi.org/10.48175/ijarsct-18035>

- Walk, O. J. B. D., Nieva, H. R., Lee, S., & Elhadad, N. (2022). A scoping review of ethics considerations in clinical *Natural Language Processing*. JAMIA Open. <https://doi.org/10.1093/jamiaopen/ooac039>
- Wu, H., Wang, M., Wu, J., Francis, F., Chang, Y., Shavick, A., Dong, H., Poon, M. T. C., Fitzpatrick, N., Levine, A. P., Slater, L. T., Handy, Á., Karwath, A., Gkoutos, G. V., Chelala, C., Shah, A. D., Stewart, R., Collier, N., Alex, B., ... Dobson, R. (2022). A survey on clinical *Natural Language Processing* in the United Kingdom from 2007 to 2022. Nature Portfolio. <https://doi.org/10.1038/s41746-022-00730-6>
- Yacouby, R., & Axman, D. (2020). Probabilistic Extension of *Precision, Recall, and F1 Score* for More Thorough Evaluation of Classification Models. EVAL4NLP. <https://doi.org/10.18653/v1/2020.eval4NLP-1.9>
- Yang, H. (2024). *Advanced Word Embeddings and Transformer-Based Feature Extraction in Natural Language Processing*. Knowledge-Based Systems, 298, 111234. DOI: <https://doi.org/10.1016/j.knosys.2024.111234>
- Yang, Z. (2024). *Deep Learning Applications in Natural Language Processing and Optimization Strategies*. Journal of Modern Education and Culture.
- Yao, J., Alabousi, A., & Mironov, O. (2024). Evaluation of a BERT *Natural Language Processing* Model for Automating CT and MRI Triage and Protocol Selection. Canadian Association of

Radiologists Journal = Journal l'Association Canadienne
Des Radiologistes.

<https://doi.org/10.1177/08465371241255895>

Yusuf, M., & Mulyana, D. I. (2025). Implementation of AI-Based *Natural Language Processing (NLP)* for Automatic Meeting Minutes and Summarization Using Voice-to-Text on Mobile Applications. *International Journal Software Engineering and Computer Science (IJSECS)*.
<https://doi.org/10.35870/ijsecs.v5i3.5459>

Zhang, X., & Li, B. (2024). A Review of the Limitations of Language Models in *NLP*. *Applied and Computational Engineering*.
<https://doi.org/10.54254/2755-2721/97/20241301>

Zhao, Y. (2025). From RNNs to BERT: A Review of Neural Models for Sequence Learning. *Theoretical and Natural Science*.
<https://doi.org/10.54254/2753-8818/2025.dl26846>

Zheng, H., Xu, K., Zhou, H., Wang, Y., & Su, G. (2024). Medication Recommendation System Based on *Natural Language Processing* for Patient Emotion Analysis. *Academic Journal of Science and Technology*.
<https://doi.org/10.54097/v160aa61>

Zheng, H., Xu, K., Zhou, H., Wang, Y., & Su, G. (2024). Medication Recommendation System Based on *Natural Language Processing* for Patient Emotion Analysis. *Academic Journal of Science and Technology*.
<https://doi.org/10.54097/v160aa61>

Buku referensi berjudul **Natural Language Processing dalam Sistem Informasi dan Kecerdasan** membahas teknologi yang memungkinkan komputer memahami, mengolah, dan menghasilkan bahasa manusia secara lebih cerdas. Perkembangan teknologi ini telah menghadirkan berbagai aplikasi yang digunakan dalam kehidupan sehari-hari, seperti penerjemah otomatis, chatbot, pencarian informasi, hingga asisten virtual.

Disajikan dengan bahasa yang mudah dipahami, buku ini mengulas konsep, penerapan, serta perkembangan terkini Natural Language Processing dalam sistem informasi dan teknologi kecerdasan. Cocok untuk masyarakat umum yang ingin mengenal lebih dekat bagaimana komputer dapat berinteraksi dengan manusia melalui bahasa dan bagaimana teknologi ini membentuk berbagai layanan digital modern.

