

## DAFTAR PUSTAKA

- [1] D. Reinsel, J. Gantz, and J. Rydning, “The Digitization of the World From Edge to Core,” 2018.
- [2] Google Cloud, “Whats is Hadoop?”
- [3] *Proceedings of the 8th International Conference on Computer Science & Education (ICCSE 2013) : the Galadari Hotel, Sri Lanka, April 26-28, 2013, Colombo, Sri Lanka*. IEEE, 2013.
- [4] Apache, “Hadoop Docs.”
- [5] E. Sivaraman and R. Manickachezian, “High performance and fault tolerant distributed file system for big data storage and processing using hadoop,” in *Proceedings - 2014 International Conference on Intelligent Computing Applications, ICICA 2014*, Institute of Electrical and Electronics Engineers Inc., Nov. 2014, pp. 32–36. doi: 10.1109/ICICA.2014.16.
- [6] F. Jaafar, G. Nicolescu, and C. Richard, “A Systematic Approach for Privilege Escalation Prevention,” in *Proceedings - 2016 IEEE International Conference on Software Quality, Reliability and Security-Companion, QRS-C 2016*, Institute of Electrical and Electronics Engineers Inc., Sep. 2016, pp. 101–108. doi: 10.1109/QRS-C.2016.17.
- [7] P. Himler, M. Landauer, F. Skopik, and M. Wurzenberger, “Towards Detecting Anomalies in Log-Event Sequences with Deep Learning: Open Research Challenges,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Jun. 2023, pp. 71–77. doi: 10.1145/3590777.3590789.
- [8] A. Sandhu and S. Mohammaed, “Detecting Anomalies in Logs by Combining NLP features with Embedding or TF-IDF.”

- [9] J. Zhu, S. He, P. He, J. Liu, and M. R. Lyu, “Loghub: A Large Collection of System Log Datasets for AI-driven Log Analytics,” Aug. 2020, [Online]. Available: <http://arxiv.org/abs/2008.06448>
- [10] H. Guo, S. Yuan, and X. Wu, “LogBERT: Log Anomaly Detection via BERT,” in *Proceedings of the International Joint Conference on Neural Networks*, Institute of Electrical and Electronics Engineers Inc., Jul. 2021. doi: 10.1109/IJCNN52387.2021.9534113.
- [11] N. Han, S. Lu, D. Wang, M. Wang, X. Tan, and X. Wei, “SKDLog: Self-Knowledge Distillation-based CNN for Abnormal Log Detection,” in *Proceedings - 2022 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Autonomous and Trusted Vehicles, Scalable Computing and Communications, Digital Twin, Privacy Computing, Metaverse, SmartWorld/UIC/ATC/ScalCom/DigitalTwin/PriComp/Metaverse 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 796–805. doi: 10.1109/SmartWorld-UIC-ATC-ScalCom-DigitalTwin-PriComp-Metaverse56740.2022.00122.
- [12] M. Sköld, “Using NLP Techniques for Log Analysis to Recommend Activities For Troubleshooting Processes.”
- [13] P. Ryciak, K. Wasielewska, and A. Janicki, “Anomaly Detection in Log Files Using Selected Natural Language Processing Methods,” *Applied Sciences (Switzerland)*, vol. 12, no. 10, May 2022, doi: 10.3390/app12105089.
- [14] P. S. Kalaki, A. Shamel-Sendi, and B. K. E. Abbasi, “Anomaly detection on OpenStack logs based on an improved robust principal component analysis model and its projection onto column space,” *Softw Pract Exp*, vol. 53, no. 3, pp. 665–681, Mar. 2023, doi: 10.1002/spe.3164.
- [15] M. Zhang, J. Chen, J. Liu, J. Wang, R. Shi, and H. Sheng, “LogST: Log Semi-supervised Anomaly Detection Based on Sentence-BERT,” in *2022 7th International Conference on Signal and Image Processing, ICSIP 2022*,

- Institute of Electrical and Electronics Engineers Inc., 2022, pp. 356–361. doi: 10.1109/ICSIP55141.2022.9886069.
- [16] V.-H. Le and H. Zhang, “Log Parsing with Prompt-based Few-shot Learning,” Feb. 2023, [Online]. Available: <http://arxiv.org/abs/2302.07435>
  - [17] Y. Huo, Y. Su, C. Lee, and M. R. Lyu, “SemParser: A Semantic Parser for Log Analysis,” Dec. 2021, [Online]. Available: <http://arxiv.org/abs/2112.12636>
  - [18] S. A. Harjanto, M. Nurhaliza, J. Hezekiah, and T. Sagala, “Optimalisasi Deteksi Anomali Untuk Pemfilteran Log dan Integrasi Dengan SIEM Menggunakan Machine Learning,” *Optimalisasi Deteksi Anomali*, vol. 2, no. 7, pp. 266–275, 2024, doi: 10.5281/zenodo.12562358.
  - [19] P. Baxendale and E. F. Codd, “Information Retrieval A Relational Model of Data for Large Shared Data Banks.”
  - [20] “Application Delivery Strategies,” 2001.
  - [21] “The Data Warehouse Toolkit.”
  - [22] . Ieee, *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies*. IEEE, 2010.
  - [23] I. Goodfellow, Y. Bengio, and A. Courville, “Deep Learning.”
  - [24] R. S. Sutton and A. G. Barto, “Reinforcement Learning: An Introduction Second edition, in progress.”
  - [25] D. Jurafsky and J. H. Martin, “Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models Third Edition draft Summary of Contents.”
  - [26] C. D. Manning and H. Schütze, “Foundations of Statistical Natural Language Processing E0123734.”

- [27] A. Vaswani *et al.*, “Attention Is All You Need,” Jun. 2017, [Online]. Available: <http://arxiv.org/abs/1706.03762>
- [28] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” [Online]. Available: <https://github.com/tensorflow/tensor2tensor>
- [29] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, “DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter,” Oct. 2019, [Online]. Available: <http://arxiv.org/abs/1910.01108>
- [30] W. Wang, F. Wei, L. Dong, H. Bao, N. Yang, and M. Zhou, “MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers,” Feb. 2020, [Online]. Available: <http://arxiv.org/abs/2002.10957>
- [31] K. Clark, M.-T. Luong, Q. V. Le, and C. D. Manning, “ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators,” Mar. 2020, [Online]. Available: <http://arxiv.org/abs/2003.10555>
- [32] P. Micikevicius *et al.*, “Mixed Precision Training,” Oct. 2017, [Online]. Available: <http://arxiv.org/abs/1710.03740>