

DAFTAR PUSTAKA

- Bachtiar, F. A., Fajar, A. N., & Maharani, A. (2018). Data Cleansing pada Data Rumah Sakit. *Jurnal Informatika Universitas Islam Indonesia*
<https://dspace.uui.ac.id/handle/123456789/10906>
- Fadlilah, R. A. (2021). Analisis dan Perancangan Modul Data Cleansing Secara Generic Menggunakan Open Source Tools (Studi Kasus: Data NIE Obat dan Makanan). *Telkom University Open Library*
<https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/13562>
- Saputra, A. Y. P. A., & Adi, K. (2023). Data Cleaning dalam Big Data: Review. *Jurnal Teknik Informatika dan Sistem Informas*
https://www.researchgate.net/publication/376758940_Data_Cleaning_dalam_Big_Data_Review
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques (3rd ed.)*. Morgan Kaufmann. (Bab tentang Data Preprocessing sering tersedia di repositori universitas)
<https://www.sciencedirect.com/science/article/pii/B9780123814791000034>
- Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing Techniques for Text Mining - An Overview. *International Journal of Computer Science & Communication Networks (IJCSCN)*, 5(1), 7-16
https://www.researchgate.net/publication/283355546_Preprocessing_Techniques_for_Text_Mining_-_An_Overview

- Cohen, W. W., Ravikumar, P., & Fienberg, S. E. (2003). A Comparison of String Distance Metrics for Name-Matching Tasks. *In Proceedings of IJCAI-03* (pp. 73-78).
<https://www.cs.cmu.edu/~wcohen/postscript/ijcai-ws-2003.pdf>
- Gomaa, W. H., & Fahmy, A. A. (2013). A Survey of Text Similarity Approaches. *International Journal of Computer Applications*, 68(13), 13-18
https://www.researchgate.net/publication/236666134_A_Survey_of_Text_Similarity_Approaches
- Mount, D. W. (2004). *Bioinformatics: Sequence and Genome Analysis* (2nd ed.). Cold Spring Harbor Laboratory Press
<https://books.google.co.id/books/about/Bioinformatics.html?id=5alrQgAACAAJ>
- Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3), 443-453.
<https://www.sciencedirect.com/science/article/abs/pii/002283670900574>
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms* (3rd ed.). MIT Press
[https://sd.blackball.lv/library/Introduction_to_Algorithms_Third_Edition_\(2009\).pdf](https://sd.blackball.lv/library/Introduction_to_Algorithms_Third_Edition_(2009).pdf)
- Eddy, S. R. (2004). What is dynamic programming?. *Nature Biotechnology*, 22(7), 909–910.
<https://www.nature.com/articles/nbt0704-909.pdf>
- McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython* (2nd ed.). O'Reilly Media

<https://www.oreilly.com/library/view/python-for-data/9781491957653/>

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace
<https://docs.python.org/3/reference/index.html>