

## DAFTAR PUSTAKA

- Alexandropoulos, S. A. N., Kotsiantis, S. B., & Vrahatis, M. N. (2019). Data preprocessing in predictive data mining. *Knowledge Engineering Review*, 34. <https://doi.org/10.1017/S026988891800036X>
- Arfianty. (2018). Analisis Penerapan Electronic Data Processing (EDP) System dalam Pengolahan Data Penjualan dan Penerimaan Kas Pada PT Hadji Kalla Cabang Parepare. In *Jurnal Ekonomi Pembangunan* (Vol. 4, Issue 1). <http://journal.stiem.ac.id/index.php/jurep/index>
- BASUKI, A. T., & YULIADI, I. (2014). *ELECTRONIC DATA PROCESSING (SPSS 15 dan EVIEWS 7)*.
- Bird, S. (2006). *NLTK: The Natural Language Toolkit*.
- Chuzaimah Zulkifli, U., & Hulliyyatus Suadaa, L. (2019). *Pengembangan Modul PreprocessingTeks untuk Kasus Formalisasi dan Pengecekan Ejaan Bahasa Indonesia pada Aplikasi Web Mining Simple Solution (WMSS)* (Vol. 15, Issue 2).
- El Naqa, I., & Murphy, M. J. (2015). What is machine learning? In *Machine Learning and the City: Applications in Architecture and Urban Design* (pp. 209–216). Wiley Blackwell. [https://doi.org/10.1007/978-3-319-18305-3\\_1](https://doi.org/10.1007/978-3-319-18305-3_1)
- Famili, A., Shen, W.-M., Weber, R., & Simoudis, E. (1997). Data preprocessing and intelligent data analysis. *Intelligent Data Analysis*, 1, 3–23.
- Filcha, A., & Hayaty, M. (2019). *Implementasi Algoritma Rabin-Karp untuk Pendeteksi Plagiarisme pada Dokumen Tugas Mahasiswa (Rabin-Karp Algorithm Implementation to Detect Plagiarism on Student's Assignment Document): Vol. VII*.
- Grijzenhout, S., & Marx, M. (2011). *The Quality of the XML Web*.

- Hassler, M., & Fliedl, G. (2006). Text preparation through extended tokenization. *WIT Transactions on Information and Communication Technologies*, 37, 13–21. <https://doi.org/10.2495/DATA060021>
- Kelly, E. J. (2018). Content Analysis of Google Alerts for Cultural Heritage Institutions. *Journal of Web Librarianship*, 12(1), 28–45. <https://doi.org/10.1080/19322909.2017.1369374>
- Khder, M. A. (2021). Web scraping or web crawling: State of art, techniques, approaches and application. *International Journal of Advances in Soft Computing and Its Applications*, 13(3), 144–168. <https://doi.org/10.15849/ijasca.211128.11>
- Ladani, D. J., & Desai, N. P. (2020). *Stopword Identification and Removal Techniques on TC and IR applications: A Survey*.
- Marta, D., Leonarde Ginting, G., & Hatuaon Sihite, A. (2022). Deteksi Berita Palsu Tentang Vaksinasi Covid-19 Dengan Menggunakan Text Mining Dan Algoritma Cosine Similarity. *Nasional Teknologi Informasi Dan Komputer*, 6(1). <https://doi.org/10.30865/komik.v6i1.5738>
- Munawar, K., Sugi, M. D., & Prabhu, V. (2021). Radiology in the News: A Content Analysis of Radiology-Related Information Retrieved From Google Alerts. *Current Problems in Diagnostic Radiology*, 50(6), 825–830. <https://doi.org/10.1067/j.cpradiol.2020.09.010>
- Nofiyani, N., & Wulandari, W. (2022). Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(3), 1621. <https://doi.org/10.30865/mib.v6i3.4332>
- Olston, C., & Najork, M. (2010). Web crawling. *Foundations and Trends in Information Retrieval*, 4(3), 175–246. <https://doi.org/10.1561/15000000017>
- Rahm, E., & Do, H. H. (2000). *Data Cleaning: Problems and Current Approaches*. <http://list.research.microsoft.com/scripts/lyris.pl?enter=debull>.

- Rosid, M. A., Fitriani, A. S., Astutik, I. R. I., Mulloh, N. I., & Gozali, H. A. (2020). Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi. *IOP Conference Series: Materials Science and Engineering*, 874(1). <https://doi.org/10.1088/1757-899X/874/1/012017>
- Sholih 'afif, M., Muzakir, M., Al, M. I., & Al Awalaien, G. (2021). TEXT MINING UNTUK MENGLASIFIKASI JUDUL BERITA ONLINE STUDI KASUS RADAR BANJARMASIN MENGGUNAKAN METODE NAÏVE BAYES. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 08(2).
- Titarsole, V., Mamgoy, G. B., & Kalalo, M. Y. B. (2019). ANALISIS DAMPAK ELECTRONIC DATA PROCESSING (EDP) TERHADAP PELAKSANAAN AUDIT INTERNAL PADA PT. BPR BANK PRISMA DANA MANADO IMPACT ANALYSIS OF ELECTRONIC DATA PROCESSING (EDP) ON THE IMPLEMENTATION OF INTERNAL AUDIT AT PT. BPR BANK PRISMA DANA MANADO. *Analisis Dampak Electro..... 3179 Jurnal EMBA*, 7(5), 3179–3188.
- Vijayarani, S., Ilamathi, Ms. J., & Nithya, Ms. (2015). *Preprocessing Techniques for Text Mining-An Overview*.
- Widodo, A. P. (2003). *Transformasi Dokumen XML*. 6(3), 128–136.
- Willett, P. (2006). The Porter stemming algorithm: Then and now. *Program*, 40(3), 219–223. <https://doi.org/10.1108/00330330610681295>