

DAFTAR PUSTAKA

- . 2016. *Data reveals the best time of year, day of the week, and time of day for finding a new job*. Diambil dari <https://careers.workopolis.com/advice/the-best-time-of-year-day-of-the-week-for-finding-a-new-job/>. Diakses pada 6 September 2020 pukul 16.20.
- Adriani, M., Asian, J., Nazief, B., Tahaghoghi, S. M., & Williams, H. E. (2007). Stemming Indonesian: A confix-stripping approach. *ACM Transactions on Asian Language Information Processing (TALIP)*, 6(4), 1-33. <https://doi.org/10.1145/1316457.1316459>
- Barbarasa, E., Barrett, J., & Goldin, N. (2017). che. Diambil dari https://economicgraph.linkedin.com/content/dam/me/economicgraph/en-us/download/Skills_Gap_or_Signalling_Gap.pdf. Diakses pada 21 Juli 2020 pukul 09.40.
- Barcaroli, G., Nurra, A., Scarnò, M., & Summa, D. (2014). Use of web scraping and text mining techniques in the Istat survey on “Information and Communication Technology in enterprises”. In *Proceedings of quality conference* (hal. 33-38). Diambil dari https://www.researchgate.net/publication/263921951_Use_of_web_scraping_and_text_mining_techniques_in_the_Istat_survey_on_Information_and_Communication_Technology_in_enterprises. Diakses pada 7 Juli 2020 pukul 11.25.
- Bilgin, Metin & Köktaş, Haldun. (2019). Sentiment Analysis with Term Weighting and Word Vectors. *International Arab Journal of Information Technology*. 5. 953-959. Diambil dari <http://paper.researchbib.com/view/paper/223901>. Diakses pada 23 Juli 2020 pukul 13.07.
- C.Ramasubramanian & R.Ramya.(2013). Effective Pre-Processing Activities in Text Mining using Improved Porter's Stemming Algorithm. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(12). Diambil dari <https://ijarccce.com/recent-issue-december-2013/>
- Canini, K., Shi, L., & Griffiths, T. (2009). Online inference of topics with latent Dirichlet allocation. *Artificial Intelligence and Statistics* (hal. 65-72). Diambil dari <https://papers.nips.cc/paper/1861-algorithms-for-non-negative-matrix-factorization>. Diakses pada 7 Agustus 2020 pukul 20.18.
- Casteel, C. A. (1988). Effects of chunked reading among learning disabled students: An experimental comparison of computer and traditional chunked

- passages. *Journal of Educational Technology Systems*, 17(2), 115-121.
<https://doi.org/10.2190/EBNP-6Q4W-1BJW-G5HM>
- Chen, Hsin-liang & Zhang, Yin. (2016). Educating Data Management Professionals: A Content Analysis of Job Descriptions. *The Journal of Academic Librarianship*. 43. <https://doi.org/10.1016/j.acalib.2016.11.002>.
- Clark, B. 2013. *The 2 Best Times of Year to Apply for Jobs (And When to Avoid)*. Diambil dari <https://careersidekick.com/the-two-best-and-worst-times-of-the-year-for-job-hunting/>. Diakses pada 6 September 2020 pukul 16.07.
- Cobos, C., Montealegre, C., Mejia, M.-F., Mendoza, M., & Leon, E. (2010). Web document clustering based on a new niching Memetic Algorithm, Term-Document Matrix and Bayesian Information Criterion. *IEEE Congress on Evolutionary Computation*. <https://doi.org/10.1109/cec.2010.5586016>
- Dang, Cach & De La Prieta, Fernando & Corchado Rodríguez, Juan & Moreno García, María. (2016). Framework for Retrieving Relevant Contents Related to Fashion from Online Social Network Data. https://doi.org/10.1007/978-3-319-40159-1_28.
- Duriau, Vincent & Reger, Rhonda & Pfarrer, Michael. (2007). A Content Analysis of the Content Analysis Literature in Organization Studies: Research Themes, Data Sources, and Methodological Refinements. *Organizational Research Methods*. 10. 5-34. <https://doi.org/10.1177/1094428106289252>.
- Ellström, P. E. (1998). The many meanings of occupational competence and qualification. Springer. *Key qualifications in work and education* (hal. 39-50). https://doi.org/10.1007/978-94-011-5204-4_3
- Gogna, N. (2014). Study of browser based automated test tools watir and selenium. *International Journal of Information & Education Technology IJIET*, 4(4). <https://doi.org/10.7763/IJIET.2014.V4.425>
- Guo, J., Mu, Y., Xiong, M., Liu, Y., & Gu, J. (2019). Activity feature solving based on TF-IDF for activity recognition in smart homes. *Complexity*, 2019. <https://doi.org/10.1155/2019/5245373>
- Hand, D. J. (2007). Principles of data mining. *Drug safety*, 30(7), 621-622. <https://doi.org/10.2165/00002018-200730070-00010>
- Harden, R. M., Sowden, S., & Dunn, W. R. (1984). Educational strategies in curriculum development: the SPICES model. *Medical education*, 18(4), 284-297. <https://doi.org/10.1111/j.1365-2923.1984.tb01024.x>
- He, J., Weerkamp, W., Larson, M., & de Rijke, M. (2009). An effective coherence measure to determine topical consistency in user-generated content. *International Journal on Document Analysis and Recognition (IJDAR)*, 12(3), 185-203. <https://doi.org/10.1007/s10032-009-0089-5>

- Hiemstra, D.(2000). A probabilistic justification for using tf×idf term weighting in information retrieval. *Int J Digit Libr* (Volume 3, hal. 131–139).
<https://doi.org/10.1007/s007999900025>
- Hintze, J. L., & Nelson, R. D. (1998). Violin Plots: A Box Plot-Density Trace Synergism. *The American Statistician*, 52(2), 181–184.
<https://doi.org/10.1080/00031305.1998.10480559>
- Ho, S. Y., & Frampton, K. (2010). A competency model for the information technology workforce: Implications for training and selection. *Communications of the Association for Information Systems*, 27(1), 5.
<https://doi.org/10.1.1.177.5456>
- Hotho, A., Nürnberger, A., & Paaß, G. (2005). A brief survey of text mining. *In Ldv Forum* (Vol. 20, No. 1, hal. 19-62). Diambil dari <http://publica.fraunhofer.de/documents/N-31569.html>. Diakses pada 17 Agustus 2020 pukul 09.18.
- Javed, F., McNair, M., Jacob, F., & Zhao, M. (2016). Towards a job title classification system. *arXiv preprint*. Diambil dari <https://arxiv.org/abs/1606.00917>. Diakses pada 29 Agustus 2020 pukul 20.15.
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169-15211.
<https://doi.org/10.1007/s11042-018-6894-4>
- Kannan, S., & Gurusamy, V. (2014). Preprocessing techniques for text mining. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16. Diambil dari <http://www.csjournals.com/?cat=4>. Diakses pada 20 Juni 2020 pukul 11.20.
- Kenthapadi, K., Le, B., & Venkataraman, G. (2017). Personalized job recommendation system at linkedin: Practical challenges and lessons learned. *Proceedings of the Eleventh ACM Conference on Recommender Systems* (hal. 346-347). <https://doi.org/10.1145/3109859.3109921>
- Kluemper, D.H. (2013). Social network screening: pitfalls, possibilities, and parallels in employment selection. *Advanced Series in Management* (hal. 1-21).
[https://doi.org/10.1108/S1877-6361\(2013\)0000012005](https://doi.org/10.1108/S1877-6361(2013)0000012005)
- Lau, J. H., Newman, D., & Baldwin, T. (2014). Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics* (hal. 530-539). <https://doi.org/10.3115/v1/E14-1056>
- Lee, D. D., & Seung, H. S. (2001). Algorithms for non-negative matrix factorization. *Advances in neural information processing systems* (hal. 556-562).
<https://doi.org/10.1.1.19.8636>

- Liu, Y., Loh, H. T., & Sun, A. (2009). Imbalanced text classification: A term weighting approach. *Expert systems with Applications*, 36(1), 690-701. <https://doi.org/10.1016/j.eswa.2007.10.042>
- Machali, I. (2014). Kebijakan perubahan kurikulum 2013 dalam menyongsong Indonesia emas tahun 2045. *Jurnal Pendidikan Islam*, 3(1), 71-94. <https://doi.org/10.14421/jpi.2014.31.71-94>
- Manning, C. D., Raghavan, P., & Schütze, H. (2009). Introduction to information retrieval. *Information Retrieval*. Cambridge: Cambridge university press. <https://doi.org/10.1109/LPT.2009.2020494>
- McLean, E. R., Smits, S. J., & Tanner, J. R. (1996). The importance of salary on job and career attitudes of information systems professionals. *Information & Management*, 30(6), 291-299. [https://doi.org/10.1016/S0378-7206\(96\)01059-2](https://doi.org/10.1016/S0378-7206(96)01059-2)
- Melling, L. (2019). What's in a name? Job title and working identity in professional services staff in higher education. *Perspectives: Policy and Practice in Higher Education*, 23(2-3), 48-53. <https://doi.org/10.1080/13603108.2018.1535459>
- Mihalcea, R., & Tarau, P. (2004). Textrank: Bringing order into text. *Proceedings of the 2004 conference on empirical methods in natural language processing* (hal. 404-411). Diambil dari <https://www.aclweb.org/anthology/W04-3252>. Diakses pada 17 Juni 2020 pukul 18.17.
- Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (hal. 262-272). <https://doi.org/10.5555/2145432.2145462>
- Moro, S., Pires, G., Rita, P., & Cortez, P. (2019). A text mining and topic modeling perspective of ethnic marketing research. *Journal of Business Research*, 103, 275-285. <https://doi.org/10.1016/j.jbusres.2019.01.053>
- Morris, P. (1996). *Hong Kong School Curriculum: Development, Issues and Policies* (Vol. 2). Hong Kong University Press. Diambil dari <https://hkupress.hku.hk/pro/con/903.pdf>. Diakses pada 7 Juli 2020 pukul 15.05.
- Newman, D., Lau, J. H., Grieser, K., & Baldwin, T. (2010). Automatic evaluation of topic coherence. *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics* (hal. 100-108). <https://doi.org/10.5555/1857999.1858011>
- Nikolenko, S. I., Koltcov, S., & Koltsova, O. (2017). Topic modeling for qualitative studies. *Journal of Information Science*, 43(1), 88-102. <https://doi.org/10.1177/0165551515617393>.

- Ningrum, P. K., Pansombut, T., & Ueranantasun, A. (2020). Text mining of online job advertisements to identify direct discrimination during job hunting process: A case study in Indonesia. *Plos one*, 15(6), e0233746. <https://doi.org/10.1371/journal.pone.0233746>
- Ramos, J. (2003). Using tf-idf to determine word relevance in document queries. *In Proceedings of the first instructional conference on machine learning* (Vol. 242, hal. 133-142). Diambil dari <https://www.cs.rutgers.edu/~mlittman/courses/ml03/iCML03/papers/ramos.pdf>. Diakses pada 5 Juli 2020 pukul 14.23.
- Razdan, R., Das, M., & Sohoni, A. (2013). The evolving Indonesian consumer. Diambil dari <https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/the-evolving-indonesian-consumer>. Diakses pada 5 September 2020 pukul 07.20.
- Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal of Documentation* (Vol. 60, hal. 503–520). <https://doi.org/10.1108/00220410410560582>
- Roos, D. (2007). How LinkedIn Works. Diambil dari <https://computer.howstuffworks.com/internet/social-networking/networks/linkedin.htm>. Diakses pada 5 September 2020 pukul 15.16.
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *Proceedings of the eighth ACM international conference on Web search and data mining* (hal. 399-408). <https://doi.org/10.1145/2684822.2685324>
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513–523. [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)
- Shariff, S. M. (2019). Investigating Selenium Usage Challenges and Reducing the Performance Overhead of Selenium-based Load Tests (Doctoral dissertation, hal 8-10). Diambil dari <https://qspace.library.queensu.ca/handle/1974/25991>. Diakses pada 22 Juli 2020 pukul 10.12.
- Shi, T., Kang, K., Choo, J., & Reddy, C. K. (2018). Short-text topic modeling via non-negative matrix factorization enriched with local word-context correlations. *Proceedings of the 2018 World Wide Web Conference* (hal. 1105-1114). <https://doi.org/10.1145/3178876.3186009>
- Simpson, G. B., Peterson, R. R., Casteel, M. A., & Burgess, C. (1989). Lexical and sentence context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(1), 88. <https://doi.org/10.1037/0278-7393.15.1.88>
- Skeels, M. M., & Grudin, J. (2009). When social networks cross boundaries: a case study of workplace use of facebook and linkedin. *In Proceedings of the*

- ACM 2009 international conference on Supporting group work* (hal. 95-104). <https://doi.org/10.1145/1531674.1531689>
- Snell, J., & Menaldo, N. (2016). Web scraping in an era of big data 2.0. *Bloomberg Law News*. Diambil dari <https://www.perkinscoie.com/en/news-insights/web-scraping-in-an-era-of-big-data-2-0.html>. Diakses pada 21 Juli 2020 pukul 12.15.
- Spärck Jones, K. (2004). IDF term weighting and IR research lessons. *Journal of Documentation* (Vol. 60, hal. 521–523). <https://doi.org/10.1108/00220410410560591>
- Sumbaly, R., Kreps, J., & Shah, S. (2013). The big data ecosystem at linkedin. *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data* (hal. 1125-1134). <https://doi.org/10.1145/2463676.2463707>
- Takey, S. & Carvalho, M. (2019). Competency mapping in project management: An action research study in an engineering company. *International Journal of Project Management*, 33, 784 – 796. <https://doi.org/10.1016/j.ijpro-man.2014.10.013>.
- Tong, Zhou & Zhang, Haiyi. (2016). A Text Mining Research Based on LDA Topic Modeling. *Computer Science & Information Technology* (Vol. 6, hal. 201-210). <https://doi.org/10.5121/csit.2016.60616>.
- Uysal, A. K., & Gunal, S. (2014). The impact of preprocessing on text classification. *Information Processing & Management*, 50(1), 104–112. <https://doi.org/10.1016/j.ipm.2013.08.006>
- Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16. Diambil dari <http://www.csjournals.com/?cat=4>. Diakses pada 22 Juli 2020 pukul 12.12.
- Verger, A., Altinyelken, H. K., & Novelli, M. (Eds.). (2018). *Global education policy and international development: New agendas, issues and policies* (2 ed.). Bloomsbury Publishing. Diambil dari <https://www.bloomsbury.com/us/global-education-policy-and-international-development-9781441143907/>
- Wang, Q., & K. W. Huang. (2018). Exploring the Fintech Jobs-Skills Fit of Financial and Information Technology Professionals: Evidence From LinkedIn. *Proceedings of the ICIS conference, Economics and IS*. Diambil dari <https://aisel.aisnet.org/icis2018/economics/Presentations/3/>. Diakses pada 27 Juli 2020 pukul 11.37.
- Xia, T., & Chai, Y. (2011). An improvement to TsF-IDF: Term Distribution based Term Weight Algorithm. *JSW*, 6(3), 413-420. <https://doi.org/10.1109/NSWCTC.2010.66>

- Zhao, B. (2017). Web scraping. *Encyclopedia of big data* (hal. 1-3).
https://doi.org/10.1007/978-3-319-32001-4_483-1
- Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., & Zou, W. (2015). A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC bioinformatics*, 16(S13), 8. <https://doi.org/10.1186/1471-2105-16-s13-s8>
- Zide, J., Elman, B., & Shahani-Denning, C. (2014). LinkedIn and recruitment: How profiles differ across occupations. *Employee Relations*, 36(5), 583-604. <https://doi.org/10.1108/ER-07-2013-0086>.