DAFTAR PUSTAKA

[1] D. Ishak, "Pelecehan Seksual Di Institusi Pendidikan: Sebuah Perspektif Kebijakan," *AKSELERASI J. Ilm. Nas.*, vol. 2, no. 2, p. 138, 2020, doi: 10.54783/jin.v2i2.462.

[2] Komisi Nasional Perempuan, "Kekerasan Seksual Di Lingkungan Pendidikan," *Https://Komnasperempuan.Go.Id/*, pp. 1–3, 2020, [Online]. Available: https://komnasperempuan.go.id/uploadedFiles/webOld/file/pdf_file/2020/Lembar Fakta KEKERASAN SEKSUAL DI LINGKUNGAN PENDIDIKAN (27 Oktober 2020).pdf.

[3] World Health Organization, "Understanding and addressing violence against women," *World Heal. Organ.*, pp. 2–3, 2012, doi: 10.1016/B978-0-08-097086-8.35026-7.

[4] A. Karami, C. N. White, K. Ford, S. Swan, and M. Yildiz Spinel, "Unwanted advances in higher education:Uncovering sexual harassment experiences in academia with text mining," *Inf. Process. Manag.*, vol. 57, no. 2, p. 102167, 2020, doi: 10.1016/j.ipm.2019.102167.

[5] A. Karami, S. Swan, and M. F. Moraes, "Space identification of sexual harassment reports with text mining," *Proc. Assoc. Inf. Sci. Technol.*, vol. 57, no. 1, pp. 1–10, 2020, doi: 10.1002/pra2.265.

[6] Hanadian Nurhayati-Wolff, "Market share of leading social media platforms in Indonesia as of July 2021," *statista*, 2021. https://www.statista.com/statistics/1256213/indonesia-social-media-market-share/ (accessed Apr. 20, 2022).

[7] K. Lubis, I. C. Nisa, P. D. Dalimunthe, and A. B. Perangin-angin, "Empathy Gap in Social Media Comments for Sexual Harassment Victim," *Int. J. Tradit. Mod. Humanit.*, vol. 2, no. 1, pp. 26–27, 2022, [Online].

Available: https://talenta.usu.ac.id/tmh.

[8] L. D. Utami and S. Masripah, “Comparation of Classification Algorithm on Sentiment Analysis of Online Learning Reviews and Distance Education,” *Techno Nusa Mandiri J. ...*, vol. 18, pp. 106–109, 2021, [Online]. Available: http://ejournal.nusamandiri.ac.id/index.php/techno/article/view/2715%0Ahttp://ejournal.nusamandiri.ac.id/index.php/techno/article/download/2715/880.

[9] A. Daderman and S. Rosander, “Evaluating Frameworks for Implementing Machine Learning in Signal Processing: A Comparative Study of CRISP-DM, SEMMA and KDD,” *Examensarbete Inom Tek.*, pp. 1–36, 2018.

[10] A. Nurdin, B. A. S. Aji, A. Bustamin, and Z. Abidin, “Perbandingan Kinerja Word Embedding Word2Vec , Glove ,” *J. TEKNOKOMPAK*, vol. 14, no. 2, pp. 77–78, 2020.

[11] A. Amalia, O. S. Sitompul, E. B. Nababan, and T. Mantoro, “An Efficient Text Classification Using fastText for Bahasa Indonesia Documents Classification,” *2020 Int. Conf. Data Sci. Artif. Intell. Bus. Anal. DATABIA 2020 - Proc.*, pp. 69–75, 2020, doi: 10.1109/DATABIA50434.2020.9190447.

[12] B. Rusyidi, A. Bintari, and H. Wibowo, “Pengalaman Dan Pengetahuan Tentang Pelecehan Seksual: Studi Awal Di Kalangan Mahasiswa Perguruan Tinggi (Experience and Knowledge on Sexual Harassment: a Preliminary Study Among Indonesian University Students),” *Share Soc. Work J.*, vol. 9, no. 1, p. 75, 2019, doi: 10.24198/share.v9i1.21685.

[13] C. Schröer, F. Kruse, and J. M. Gómez, “A systematic literature review on applying CRISP-DM process model,” *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.

[14] P. Bhatia, *Data mining and data warehousing : principles and practical*

techniques. Cambridge Unversity Press, 2019.

[15] T. Jo, *Text Mining: Concepts, Implementation, and Big Data Challenge*, 1st ed., vol. 45. Springer, Cham, 2019.

[16] Z. Kastrati, F. Dalipi, A. S. Imran, K. P. Nuci, and M. A. Wani, "Sentiment analysis of students' feedback with nlp and deep learning: A systematic mapping study," *Appl. Sci.*, vol. 11, no. 9, pp. 3–5, 2021, doi: 10.3390/app11093986.

[17] H. Lane, C. Howard, and H. M. Hapke, *Natural Language Processing in Action(Understanding,analyzing, and generating text with python)*. 2019.

[18] H. Ma'rifah, A. P. Wibawa, and M. I. Akbar, "Klasifikasi Artikel Ilmiah Dengan Berbagai Skenario Preprocessing," *Sains, Apl. Komputasi dan Teknol. Inf.*, vol. 2, no. 2, p. 71, 2020, doi: 10.30872/jsakti.v2i2.2681.

[19] R. Khan *et al.*, "Social Media Analysis With Ai: Sentiment Analysis Techniques for the Analysis of Twitter Covid-19 Data Journal of Critical Reviews Social Media Analysis With Ai: Sentiment Analysis Techniques for the Analysis of Twitter Covid-19 Data," *Artic. J. Crit. Rev.*, vol. 7, no. 09, pp. 2765–2766, 2020, [Online]. Available: https://www.researchgate.net/publication/343685163.

[20] L. Liu and M. T. Özsu, *Ling Liu, M. Tamer Özsu - Encyclopedia of Database Systems-Springer New York (2018).pdf*, Second Edi. Springer Science+Business Media, LLC 2009 (USA), 2018.

[21] F. Rahutomo and A. R. T. H. Ririd, "Evaluasi Daftar Stopword Bahasa Indonesia," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 1, p. 41, 2019, doi: 10.25126/jtiik.2019611226.

[22] K. Divya, B. Siddhartha, N. Niveditha, and B. Divya, "An Interpretation of Lemmatization and Stemming in Natural Language Processing," *J. Univ. Shanghai Sci. Technol.*, vol. 22, no. 10, p. 353, 2020, [Online]. Available: https://www.researchgate.net/publication/348306833.

[23] J. Abuja and J. Abuja, “a Natural Language Processing Approach To Determine the Polarity and Subjectivity of Iphone 12 Twitter Feeds Using,” pp. 0–7, 2021, doi: 10.52417/ojps.v2i2.276.

[24] A. Mandelbaum and A. Shalev, “Word Embeddings and Their Use In Sentence Classification Tasks,” 2016, [Online]. Available: http://arxiv.org/abs/1610.08229.

[25] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, “Enriching Word Vectors with Subword Information,” *Trans. Assoc. Comput. Linguist.*, vol. 5, pp. 135–146, 2017, doi: 10.1162/tacl_a_00051.

[26] L. McInnes, J. Healy, and J. Melville, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,” 2020, [Online]. Available: http://arxiv.org/abs/1802.03426.

[27] A. M. Michael W. Berry, “Supervised and unsupervised learning,” *Springer Nat. Switz. AG 2020*, no. January, pp. 4–15, 2020, doi: 10.1007/978-3-030-22475-2_1.

[28] S. Huang, C. A. I. Nianguang, P. Penzuti Pacheco, S. Narandes, Y. Wang, and X. U. Wayne, “Applications of support vector machine (SVM) learning in cancer genomics,” *Cancer Genomics and Proteomics*, vol. 15, no. 1, pp. 41–42, 2018, doi: 10.21873/cgp.20063.

[29] P. Diez, *Introduction*, Second Edi. Elsevier B.V., 2018.

[30] A. Kulkarni, D. Chong, and F. A. Batarseh, *Foundations of data imbalance and solutions for a data democracy*. Elsevier Inc., 2020.

[31] A. Karami, M. Y. Spinel, C. N. White, K. Ford, and S. Swan, “A systematic literature review of sexual harassment studies with text mining,” *Sustain.*, vol. 13, no. 12, pp. 1–24, 2021, doi: 10.3390/su13126589.

[32] K. Budiman, N. Zaatsiyah, U. Niswah, F. Muhanna, and N. Faizi, “Analysis of Sexual Harassment Tweet Sentiment on Twitter in Indonesia

using Naïve Bayes Method through National Institute of Standard and Technology Digital Forensic Acquisition Approach," *J. Adv. Inf. Syst. Technol.*, vol. 2, no. 2, pp. 21–30, 2020, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/jaist.

[33] A. A. Lutfi, A. E. Permanasari, and S. Fauziati, "Sentiment Analysis in the Sales Review of Indonesian Marketplace by Utilizing Support Vector Machine," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 4, no. 2, pp. 61–63, 2018, doi: 10.20473/jisebi.4.2.169.

[34] W. Nengsih, "Analisa Akurasi Permodelan Supervised Dan Unsupervised," *Sebatik 1410-3737*, vol. 23, no. 2, pp. 287–291, 2017, [Online]. Available: https://jurnal.wicida.ac.id/index.php/sebatik/article/view/771.

[35] K. B. Vadloori and S. M. Sanghishetty, "Exploratory and Sentiment Analysis of Netflix Data," *Int. J. Eng. Res. Technol.*, vol. 10, no. 09, pp. 214–216, 2021, [Online]. Available: https://www.ijert.org/exploratory-and-sentiment-analysis-of-netflix-data.

[36] Q. H. Nguyen *et al.*, "Influence of data splitting on performance of machine learning models in prediction of shear strength of soil," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/4832864.

[37] N. Bhanot, H. Singh, D. Sharma, H. Jain, and S. Jain, "Python vs. R: A Text Mining Approach for analyzing the Research Trends in Scopus Database," pp. 20–24, 2019, [Online]. Available: http://arxiv.org/abs/1911.08271.