

## DAFTAR PUSTAKA

- Aizawa, A. (2003). An information-theoretic perspective of tf-idf measures. *Information Processing & Management*, 39(1), 45–65.
- Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., & Ridella, S. (2012). The ‘K’ in K-fold cross validation. *20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 441–446.
- Awad, M., & Khanna, R. (2015). Support vector machines for classification. *Efficient Learning Machines*, 39–66.
- Brugha, R., & Varvasovszky, Z. (2000). Stakeholder analysis: a review. *Health Policy and Planning*, 15(3), 239–246.
- Dietterich, T. G. (1997). Machine-learning research. *AI Magazine*, 18(4), 97–97. <https://ojs.aaai.org/index.php/aimagazine/article/view/1324>
- Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25, 1–54. <https://www.jstatsoft.org/article/view/v025i05>
- Fukunaga, K., & Narendra, P. M. (1975). A branch and bound algorithm for computing k-nearest neighbors. *IEEE Transactions on Computers*, 100(7), 750–753.
- Fushiki, T. (2011). Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*, 21(2), 137–146.
- Gao, X., & Li, G. (2020). A KNN model based on manhattan distance to identify the SNARE proteins. *Ieee Access*, 8, 112922–112931. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9119343>
- Gupta, V., & Lehal, G. S. (2009). A survey of text mining techniques and applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60–76.

- Hotho, A., Nürnberger, A., & Paaß, P. (2005). A brief survey of text mining. *Ldv Forum*, 20(1), 19–62.  
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.447.4161&rep=rep1&type=pdf>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260.  
<https://www.science.org/doi/abs/10.1126/science.aaa8415>
- Khadapi, M., Riana, D., Arfian, A., & Rahmawati, E. (2020). Public Acceptance Of Pedulilindungi Application In The Acceleration Of Corona Virus (Covid-19) Handling. *Journal of Physics: Conference Series*, 1641(1), 012026.
- Krouska, A., Troussas, C., & Virvou, M. (2016). The effect of preprocessing techniques on Twitter sentiment analysis. *2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA)*, 1–5.
- Luque, A., Carrasco, A., Martín, A., & de Las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91, 216–231.
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, 9, 381–386.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113.
- Millman, K. J., & Aivazis, M. (2011). Python for scientists and engineers. *Computing in Science & Engineering*, 13(2), 9–12.
- Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine learning with oversampling and undersampling techniques: overview study and experimental results. *2020 11th International Conference on Information and Communication Systems (ICICS)*, 243–248.

- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 5(2), 697–711.
- Nurjanah, W. E., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-ISSN, 2548, 964X*.
- Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science & Engineering*, 9(3), 10–20.
- Rahutomo, F., Kitasuka, T., & Aritsugi, M. (2012). Semantic cosine similarity. *The 7th International Student Conference on Advanced Science and Technology ICAST*, 4(1), 1.
- Rajman, M., & Besançon, R. (1998). Text mining: natural language techniques and text mining applications. *Data Mining and Reverse Engineering*, 50–64. [https://link.springer.com/content/pdf/10.1007/978-0-387-35300-5\\_3.pdf](https://link.springer.com/content/pdf/10.1007/978-0-387-35300-5_3.pdf)
- Ramos, J. (2003). Using tf-idf to determine word relevance in document queries. *Proceedings of the First Instructional Conference on Machine Learning*, 242(1), 29–48.
- Shah, K., Patel, H., Sanghvi, D., & Shah, M. (2020). A comparative analysis of logistic regression, random forest and KNN models for the text classification. *Augmented Human Research*, 5(1), 1–16.
- Sudiarsa, I. W., & Wiraditya, I. G. B. (2020). Analisis Usability Pada Aplikasi Peduli Lindungi Sebagai Aplikasi Informasi Dan Tracking Covid-19 Dengan Heuristic Evaluation. *INTECOMS: Journal of Information Technology and Computer Science*, 3(2), 354–364.
- Viennot, N., Garcia, E., & Nieh, J. (2014). A measurement study of google play. *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems*, 221–233.

- 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.
- Yap, B. W., Rani, K. A., Rahman, H. A. A., Fong, S., Khairudin, Z., & Abdullah, N. N. (2014). An application of oversampling, undersampling, bagging and boosting in handling imbalanced datasets. *Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)*, 13–22.
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.
- Zhang, Z. (2016). Introduction to machine learning: k-nearest neighbors. *Annals of Translational Medicine*, 4(11). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4916348/pdf/atm-04-11-218.pdf>
- Zhong, N., & Michahelles, F. (2013). Google Play is not a long tail market: An empirical analysis of app adoption on the Google Play app market. *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, 499–504.