

## DAFTAR PUSTAKA

- Ahdiat, A. (2022). *Setelah Terjun Bebas, Saham GOTO Mulai Naik Lagi?* Katadata.
- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91–93.
- Alkubaisi, G. A. A. J. (2019). The Role of Ensemble Learning in Stock Market Classification Model Accuracy Enhancement Based on Naive Bayes Classifiers. *International Journal of Statistics and Probability*, 9(1), 36.
- Al-Radaideh, Q., Assaf, A., & Alnagi, E. (2013). Predicting stock prices using data mining techniques. *The International Arab Conference on Information Technology, December 2013*, 8.
- Al-Shamiri, A. Y. R. (2021). Artificial intelligence and pattern recognition using data mining algorithms. *International Journal of Computer Science & ...*, 21(7).
- Al-Taie, M. Z., Kadry, S., & Lucas, J. P. (2019). Online data preprocessing: A case study approach. *International Journal of Electrical and Computer Engineering*, 9(4), 2620–2626.
- Alam, S., & Yao, N. (2019). The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis. *Computational and Mathematical Organization Theory*, 25(3), 319–335.
- Alan, R. H., Salvatore, T. M., Jinsoo, P., & Sudha, R. (2004). Design science 97. *AI and Society*, 10(2), 199–217.
- Analytics, P. A. M. (1996). *The CRISP-DM Methodology*. 1–3.
- Annur, C. (2022). *Pengguna Twitter Indonesia Masuk Daftar Terbanyak di Dunia, Urutan Berapa?* Katadata.
- Ariesanto Akhmad, E. P. (2020). Data Mining Menggunakan Regresi Linear untuk Prediksi Harga Saham Perusahaan Pelayaran. *Jurnal Aplikasi Pelayaran dan Kepelabuhanan*, 10(2), 120.

- Aws. (2022). *Apa itu Analisis Sentimen?* Amazon Web Services.
- Bakheet, S., & Al-Hamadi, A. (2021). A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve bayesian classification. *Brain Sciences*, *11*(2), 1–15.
- Batista, G. E. A. P. A., Bazzan, A. L. C., & Monard, M. C. (2003). Balancing Training Data for Automated Annotation of Keywords: a Case Study. *In Proceedings of the Second Brazilian Workshop on Bioinformatics, January*, 35–43.
- Bruck, T. Der. (2022). *The F-Measure Paradox The F-Measure Paradox*. February.
- Carosia, A. E. O., Coelho, G. P., & Silva, A. E. A. (2020). Analyzing the Brazilian Financial Market through Portuguese Sentiment Analysis in Social Media. *Applied Artificial Intelligence*, *34*(1), 1–19.
- Challa, S. K., & Vamsi, B. (2022). *Attitudinal Disposition of Small Investors on Mutual Funds : an Empirical Slanton Selected Districts of Andhra Region*. February.
- Chandra, W., Suprihatin, B., & Resti, Y. (2023). Median-KNN Regressor-SMOTE-Tomek Links for Handling Missing and Imbalanced Data in Air Quality Prediction. *Symmetry*, *15*(4).
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). Crisp-Dm. *SPSS inc*, *78*, 1–78.
- Chen, X., Xie, H., Cheng, G., Poon, L. K. M., Leng, M., & Wang, F. L. (2020). Trends and features of the applications of natural language processing techniques for clinical trials text analysis. *Applied Sciences (Switzerland)*, *10*(6), 1–36.
- Chowdhary, K. R. (2020). Fundamentals of artificial intelligence. *In Fundamentals of Artificial Intelligence*.
- Deddy, D. (2020). Implementasi Web Crawling untuk Pencarian Harga Sparepart

- Pada PT Asuransi Sinar Mas. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 7(3), 416–428.
- Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier. *International Journal of Information Engineering and Electronic Business*, 8(4), 54–62.
- Eisenstein, J. (2019). *Introduction to Natural Language Processing (Adaptive c)*. The MIT Press.
- Gautam, A. K., & Bansal, A. (2022). Performance Analysis of Supervised Machine Learning Techniques for Cyberstalking Detection in Social Media. *Journal of Theoretical and Applied Information Technology*, 100(2), 449–461.
- GoTo. (2020). *GoTo Company*. Goto Company.
- GoTo. (2021). *Sekilas Tentang GOTO*. Goto Company.
- Hairani, H., Anggrawan, A., & Priyanto, D. (2023). Improvement Performance of the Random Forest Method on Unbalanced Diabetes Data Classification Using Smote-Tomek Link. *JOIV: International Journal on Informatics Visualization*, 7(1), 258.
- Hassani, A., Iranmanesh, A., & Mansouri, N. (2021). Text mining using nonnegative matrix factorization and latent semantic analysis. *Neural Computing and Applications*, 33(20), 13745–13766.
- Humaira, A., & Tejomurti, K. (2022). Analysis of Consumer Personal Data Protection for Data Integration: the Gojek-Tokopedia Case. *Indonesian Journal of Law and Policy Studies*, 3(1), 20–30.
- Jifan, C., & Thalhah, M. (2023). Retracted: Audit Data Analysis and Application Based on Correlation Analysis Algorithm. *Computational and mathematical methods in medicine*, 2023, 9821320.
- Kewsuwun, N., & Kajornkasirat, S. (2022). A sentiment analysis model of

- agritech startup on Facebook comments using naive Bayes classifier. *International Journal of Electrical and Computer Engineering*, 12(3), 2829–2838.
- Khan, I. U., Khan, A., Khan, W., Su'ud, M. M., Alam, M. M., Subhan, F., & Asghar, M. Z. (2022). A Review of Urdu Sentiment Analysis With Multilingual Perspective: A Case of Urdu and Roman Urdu Language. *Computers*, 11(1), 1–29.
- Khattak, A. M., Ullah, H., Khalid, H. A., Habib, A., Asghar, M. Z., & Kundi, F. M. (2019). Stock market trend prediction using supervised learning. *ACM International Conference Proceeding Series*, ii, 85–91.
- Kimmons, R., Rosenberg, J., & Allman, B. (2021). Trends in Educational Technology: What Facebook, Twitter, and Scopus Can Tell us about Current Research and Practice. *TechTrends*, 65(2), 125–136.
- Kvam, P., Vidakovic, B., & Kim, S.-J. (2023). *Nonparametric Statistics with Applications to Science and Engineering with R* (J. Wiley & Sons (ed.); Second Edi). John Wiley & Sons, Inc.
- Linawati, S., Safitri, R. A., Alfiyan, A. R., Pangesti, W. E., Winarto, M. N., Bina, U., Informatika, S., Bayes, N., & Machine, S. V. (2020). Perbandingan Algoritma Klasifikasi Naive Bayes Dan Svm. *Jurnal Swabumi*, 8(1), 71–75.
- Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing, Second Edition*, 627–666.
- Liu, B. (2012). Sentiment analysis and opinion mining. In G. Hirst (Ed.), *Synthesis Lectures on Human Language Technologies* (Vol. 5, Nomor 1). Springer series.
- Mani Sekhar, S. R., Siddesh, G. M., Manvi, S. S., & Srinivasa, K. G. (2019). Optimized focused Web Crawler with Natural Language Processing based relevance measure in bioinformatics web sources. *Cybernetics and Information Technologies*, 19(2), 146–158.
- MetroTV. (2022). *Harga Saham GoTo Turun 6,6% Hari ini*. Metro TV.

- Mihalcea, R. (2007). The text mining handbook: advanced approaches in analyzing unstructured data. *Choice Reviews Online*, 44(10), 44-5684-44-5684.
- Mitra, A. (2020). Sentiment Analysis Using Machine Learning Approaches (Lexicon based on movie review dataset). *Journal of Ubiquitous Computing and Communication Technologies*, 2(3), 145–152.
- Mudinas, A., Zhang, D., & Levene, M. (2019). *Market Trend Prediction using Sentiment Analysis: Lessons Learned and Paths Forward*.
- Mustaqim, T., Umam, K., & Muslim, M. A. (2020). Twitter text mining for sentiment analysis on government's response to forest fires with vader lexicon polarity detection and k-nearest neighbor algorithm. *Journal of Physics: Conference Series*, 1567(3).
- Nurfadhila, B., & Girsang, A. S. (2023). Identifying Indication of Depression of Twitter User in Indonesia Using Text Mining. *International Journal of Intelligent Systems and Applications in Engineering*, 11(2), 523–530.
- Novendri, R., Callista, A. S., Pratama, D. N., & Puspita, C. E. (2020). Sentiment Analysis of YouTube Movie Trailer Comments Using Naïve Bayes. *Bulletin of Computer Science and Electrical Engineering*, 1(1), 26–32.
- Nurrohmat, M. A., & Azhari, S. (2019). Sentiment Analysis of Novel Review Using Long Short-Term Memory Method. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 13(3), 209.
- Oliveira, D. N. de, & Merschmann, L. H. de C. (2021). Joint evaluation of preprocessing tasks with classifiers for sentiment analysis in Brazilian Portuguese language. *Multimedia Tools and Applications*, 80(10), 15391–15412.
- Oluwaseun, A., & Chaubey, M. S. (2019). Data Mining Classification Techniques on the analysis of student performance. *Global Scientific Journal*, 7(April), 79–95.
- Padhanarath, P., Aunhathaweep, Y., & Kiattisin, S. (2019). Sentiment analysis

- and relationship between social media and stock market: Pantip.com and SET. *IOP Conference Series: Materials Science and Engineering*, 620(1).
- Padmanayana, Varsha, & Bhavya, K. (2021). Stock market prediction using sentiment analysis. *International Journal of Scientific Research in Computer Science*.
- Pratama, A. H., & Hayaty, M. (2023). *Performance of Lexical Resource and Manual Labeling on Long Short-Term Memory Model for Text Classification*. 9(1), 74–84.
- Putra, D., & Wibowo, A. (2020). Prediksi Keputusan Minat Penjurusan Siswa SMA Yadika 5 Menggunakan Algoritma Naïve Bayes. *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, 2, 84–92.
- Rani, R., & Lobiyal, D. K. (2022). Performance evaluation of text-mining models with Hindi stopwords lists. *Journal of King Saud University - Computer and Information Sciences*, 34(6), 2771–2786.
- Renault, T. (2020). Sentiment analysis and machine learning in finance: a comparison of methods and models on one million messages. *Digital Finance*, 2(1–2), 1–13.
- Rizaty, M. (2022). *Pengguna Twitter di Indonesia Capai 18,45 Juta pada 2022*. Data Indonesia.
- Rosenoff, D., Chatelain, E., Bandepalli, V., Kottapuzhackal, S., & Karunakaran, V. (2022). *Using Confusion Matrix Statistics to determine HRT Study Quality*.
- Rukhsar, L., Haider Bangyal, W., Nisar, K., & Nisar, S. (2022). Prediction of Insurance Fraud Detection using Machine Learning Algorithms. *Mehran University Research Journal of Engineering and Technology*, 41(1), 33–40.
- Rybachak, Z., & Basystiuk, O. (2017). Analysis of methods and means of text mining. *ECONTECHMOD. An International Quarterly Journal*, 6(2), 73–78.

- Sabti, A. A., Rashid, S. M., & Hummadi, A. S. (2019). Interrelationships between writing anxiety dimensions and writing goal orientation among Iraqi EFL undergraduates. *International Journal of Instruction*, 12(4), 529–544.
- Salma, A., & Silfianti, W. (2021). Sentiment Analysis of User Review on COVID-19 Information Applications Using Naïve Bayes Classifier, Support Vector Machine, and K-Nearest Neighbors. *International Research Journal of Advanced Engineering and Science*, 6(4), 158–162.
- Santoso, V. I., Virginia, G., & Lukito, Y. (2017). Penerapan Sentimen Analisis Pada Hasil Evaluasi Dosen Dengan Metode SVM. *Jurnal Transformatika*, 14(2), 72.
- Schober, P., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia*, 126(5), 1763–1768.
- SimilarWeb. (2022). *Top Websites Ranking*. Similar Web.
- Sjarif, N. N. A., Azmi, N. F. M., Chuprat, S., Sarkan, H. M., Yahya, Y., & Sam, S. M. (2019). SMS spam message detection using term frequency-inverse document frequency and random forest algorithm. *Procedia Computer Science*, 161, 509–515.
- Sreedharan, R., & Praveen Kumar, A. (2021). Prediction of Stock Market Prices using Machine Learning. *Scrs Conference Proceedings on Intelligent Systems, May*, 339–347.
- Sueno, H., Medina, R., & Gerardo, B. D. (2020). Multi-class Document Classification using Support Vector Machine (SVM) Based on Improved Naïve Bayes Vectorization Technique. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(3), 3937–3944.
- Swana, E. F., Doorsamy, W., & Bokoro, P. (2022). Tomek Link and SMOTE Approaches for Machine Fault Classification with an Imbalanced Dataset. *Sensors*, 22(9).
- Syaifudin, Y. W., & Puspitasari, D. (2017). Twitter Data Mining for Sentiment Analysis on Peoples Feedback Against Government Public Policy.

- MATTER: International Journal of Science and Technology*, 3(1), 110–122.
- ULUDAĞ, O., & GÜRSOY, A. (2020). On the Financial Situation Analysis with KNN and Naive Bayes Classification Algorithms. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 10(4), 2881–2888.
- Wibowo, A. (2022). *Mengenal Aplikasi Yahoo Finance Indonesia*. PT Kredit Pintar Indonesia.
- Widiyaningsih, S. D., & Pertiwi, A. (2020). Analysis of Ovo Application Sentiment Using Lexicon Based Method and K-Nearest Neighbor. *Jurnal Ilmiah Ekonomi Bisnis*, 25(1), 14–28.
- Wikanto, A. (2022). *Harga Saham GOTO Sentuh Level Terendah, Hari Ini (1/12) Pilih Beli Atau Jual?* Kontan.co.id.
- Wijesinghe, C. (2020). *Tracking the Corona Virus Using Sentiment Analysis on Twitter Data: Application to Predict Stock Market Movements*. 139.
- Yahoo. (2022). *Yahoo Finance*. Yahoo finance.
- Zhang, B., & Kong, D. (2020). Dynamic estimation model of insurance product recommendation based on Naive Bayesian model. *ACM International Conference Proceeding Series*, 219–224.