

DAFTAR PUSTAKA

- Achsan, H. T. Y., Kurniawan, D., Purnama, D. G., Barcah, Q. K. D., & Astoria, Y. Y. (2022). Application of Natural Language Processing Using Cosine-Similarity Algorithm in Making Chatbot Information on the New Capital City of the Republic of Indonesia. *IWBIS 2022 - 7th International Workshop on Big Data and Information Security, Proceedings*, 1–6. <https://doi.org/10.1109/IWBIS56557.2022.9924902>
- Adinugroho, S., & Sari, Y. A. (2018). *Implementasi Data Mining Menggunakan Weka*. Universitas Brawijaya Press.
- Ardianto, A. S., & Adi, S. (2022). The Best Problem Transformation Method in Multi-Label Classification Text for Thesis Abstract. *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 326–329. <https://doi.org/10.1109/icitisee57756.2022.10057824>
- Ashari, H., Arifianto, D., Azizah, H., & Faruq, A. (2020). Perbandingan Kinerja Algoritma Multinomial Naive Bayes (MNB, Multivariate Bernoulli dan Rocchio Algoritma) Dalam Klasifikasi Konten Berita Hoax Berbahasa Indonesia Pada Media Sosial. *Http://Repository.Unmuhjember.Ac.Id*, 1–12.
- Balas, Valentina E and Kumar, Raghvendra and Srivastava, & Rajshree and others. (2020). *Recent Trends and Advances in Artificial Intelligence and Internet of Things*.
- Berger, M. J. (2014). Large Scale Multi-label Text Classification with Semantic Word Vectors. *Technical Report*, 1–8.
- Cahyani, L., & Arif, M. (2022). Text Mining untuk Pengelompokan Skripsi di Prodi Pendidikan Informatika Universitas Trunojoyo Madura. *Jurnal Ilmiah Edutic : Pendidikan Dan Informatika*, 8(2), 97–108. <https://doi.org/10.21107/edutic.v8i2.13020>
- Choirulfikri, M. R., Lhaksamana, K. M., & Faraby, S. Al. (2022). A Multi-Label Classification of Al-Quran Verses Using Ensemble Method and Naïve Bayes. *Building of Informatics, Technology and Science (BITS)*, 3(4), 473–479. <https://doi.org/10.47065/bits.v3i4.1287>
- Dewi, F. K. S. (2021). Klasifikasi Berita Menggunakan Metode Multinomial Naïve Bayes. *SCAN - Jurnal Teknologi Informasi Dan Komunikasi*, 16(3), 1–8. <https://doi.org/10.33005/scan.v16i3.2870>
- Faisal, M. R., & Nugrahadi, D. T. (2017). *Belajar Data Science: Klasifikasi dengan Bahasa Pemrograman R*. M Reza Faisal.

- Firdaus, A., & Firdaus, W. I. (2021). Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan). *Jurnal JUPITER*, 13(1), 66.
- Gibaja, E., & Ventura, S. (2015). A tutorial on multilabel learning. *ACM Computing Surveys*, 47(3). <https://doi.org/10.1145/2716262>
- Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications - Volume 1, No. 1, August 2009 - JETWI. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60–76.
- Hadaina, F., & Budiyo, U. (2022). Implementasi Metode Multinomial Naïve Bayes Untuk Sentiment Analysis Terhadap Data Ulasan Produk Clearn Pada Google Play Store Implementation Of Multinomial Naive Bayes Method For Sentiment Analysis Of Clearn Product Review Data On Google Play Store. *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) Jakarta-Indonesia, September*, 660–666.
- Hairani, H., Anggrawan, A., Wathan, A. I., Latif, K. A., Marzuki, K., & Zulfikri, M. (2021). The Abstract of Thesis Classifier by Using Naive Bayes Method. *Proceedings - 2021 International Conference on Software Engineering and Computer Systems and 4th International Conference on Computational Science and Information Management, ICSECS-ICOCSIM 2021*, 312–315. <https://doi.org/10.1109/ICSECS52883.2021.00063>
- Hanafi, A., Adiwijaya, A., & Astuti, W. (2020). Klasifikasi Multi Label pada Hadis Bukhari Terjemahan Bahasa Indonesia Menggunakan Mutual Information dan k-Nearest Neighbor. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 9(3), 357–364. <https://doi.org/10.32736/sisfokom.v9i3.980>
- Isnaini, N., Adiwijaya, Mubarak, M. S., & Bakar, M. Y. A. (2019). A multi-label classification on topics of Indonesian news using K-Nearest Neighbor. *Journal of Physics: Conference Series*, 1192(1). <https://doi.org/10.1088/1742-6596/1192/1/012027>
- Joachims, T. (1996). *A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization*. *Computer Science Technical Report CMU-CS-96-118*. March.
- Jurafsky, D., & Martin, J. H. (2024). *Speech and Language Processing*. <https://doi.org/10.4159/9780674039315-001>
- Kadhim, A. I. (2018). An Evaluation of Preprocessing Techniques for Text Classification. *International Journal of Computer Science and Information Security*, 16(6), 22–32.
- Kalt, T. (1996). *A new probabilistic model of text classification and retrieval*.



- Karunia, S. A., Saptono, R., & Anggrainingsih, R. (2017). Online News Classification Using Naive Bayes Classifier with Mutual Information for Feature Selection. *Jurnal Ilmiah Teknologi Dan Informasi*, 6(1), 10–15.
- Kibriya, A. M., Frank, E., Pfahringer, B., & Holmes, G. (2004). Multinomial naive bayes for text categorization revisited. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 3339, 488–499. https://doi.org/10.1007/978-3-540-30549-1_43
- Kim, H. C., Park, J. H., Kim, D. W., & Lee, J. (2020). Multilabel naïve Bayes classification considering label dependence. *Pattern Recognition Letters*, 136, 279–285. <https://doi.org/10.1016/j.patrec.2020.06.021>
- Larsen, P. O., & von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by science citation index. *Scientometrics*, 84(3), 575–603. <https://doi.org/10.1007/s11192-010-0202-z>
- Lewis, D. D. (1995). A Sequential Algorithm for Training Text Classifiers: Corrigendum and Additional Data. *ACM SIGIR Forum*, 29(2), 13–19. <https://doi.org/10.1145/219587.219592>
- Li, H., & Yamanishi, K. (1997). Document classification using a finite mixture model. *Proceedings of the Annual Meeting of the Association for Computational Linguistics, 1997-July*, 39–47. <https://doi.org/10.3115/979617.979623>
- Liu, H., Chen, G., Li, P., Zhao, P., & Wu, X. (2021). Multi-label text classification via joint learning from label embedding and label correlation. *Neurocomputing*, 460, 385–398. <https://doi.org/10.1016/j.neucom.2021.07.031>
- Mallamma, V., & Hanumanthappa, M. (2014). Semantical and Syntactical Analysis of NLP. *International Journal of Computer Science and Information Technologies*, 5(3), 3236–3238.
- Manueke, M., & Sengkey, R. (2022). *Implementation of Multi-Label Classification to Determine Scientific Articles Keyword in Journals*. 1–7.
- McCallum, A., & Nigam, K. (1998a). A Comparison of Event Models for Naive Bayes Text Classification. *AAAI/ICML-98 Workshop on Learning for Text Categorization*, 41–48. <https://doi.org/10.1.1.46.1529>
- McCallum, A., & Nigam, K. (1998b). A Comparison of Event Models for Naive Bayes Text Classification. *AAAI/ICML-98 Workshop on Learning for Text Categorization*, 41–48. <https://doi.org/10.1.1.46.1529>
- Mubarok, Huda, M. S. and, Saiful, N., & Others. (2018). A Multi-label Classification on Topics of Quranic Verses in English Translation Using Tree Augmented Naïve Bayes. *Journal of Physics: Conference Series*, 1192(1),

103–106. <https://doi.org/10.1088/1742-6596/1192/1/012026>

- Mustafa, G., Usman, M., Yu, L., afzal, M. T., Sulaiman, M., & Shahid, A. (2021). Multi-label classification of research articles using Word2Vec and identification of similarity threshold. *Scientific Reports*, 11(1), 1–20. <https://doi.org/10.1038/s41598-021-01460-7>
- Myint Maw, T. M., Clenton, J., & Higginbotham, G. (2022). Investigating whether a flemma count is a more distinctive measurement of lexical diversity. *Assessing Writing*, 53(October 2021), 100640. <https://doi.org/10.1016/j.asw.2022.100640>
- Nugraha, Fikri Aldi and Harani, Nisa Hanum and Habibi, & Roni. (2020). *Analisis Sentimen Terhadap Pembatasan Sosial Menggunakan Deep Learning*. Kreatif.
- Pane, R. A., Mubarak, M. S., & Adiwijaya. (2018). Klasifikasi Multi-Label Pada Topik Ayat Al-Quran Terjemahan Bahasa Inggris Menggunakan Multinomial Naive Bayes. *E-Proceeding of Engineering*, 5(1), 1551–1555.
- Primartha, R. (2021). *Algoritma Machine Learning*. Informatika Bandung.
- Puttipornchai, C., Chanyachatchawan, S., & Tuaycharoen, N. (2022). Multi-Label Classification for Articles in Thai Journal Database from Article's Abstract. *2022 19th International Joint Conference on Computer Science and Software Engineering, JCSSE 2022*, 1–6. <https://doi.org/10.1109/JCSSE54890.2022.9836270>
- Qorib, M., Oladunni, T., Denis, M., Ososanya, E., & Cota, P. (2023). Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset. *Expert Systems with Applications*, 212(September 2022), 118715. <https://doi.org/10.1016/j.eswa.2022.118715>
- Rahayu, S., Fatayat, & Salambue, R. (2020). *Analisis Penentuan Dosen Pembimbing Skripsi Mahasiswa Menggunakan Naive Bayes Classifier*. 1–10.
- Rahman, A., Wiranto, & Doewes, A. (2017). Online News Classification Using Multinomial Naive Bayes. *Jurnal Ilmiah Teknologi Dan Informasi*, 6(1), 32–38.
- Rinaldi, R., Goejantoro, R., & Syaripuddin. (2021). Penerapan Metode Klasifikasi Multinomial Naive Bayes (Studi Kasus: PT Prudential Life Samarinda Tahun 2019). *Eksponensial*, 12(2), 111–118.
- Ristanti, P. Y., Wibawa, A. P., & Pujianto, U. (2019). Cosine Similarity for Title and Abstract of Economic Journal Classification. *Proceeding - 2019 5th International Conference on Science in Information Technology: Embracing Industry 4.0: Towards Innovation in Cyber Physical System, ICSITech 2019*, 123–127. <https://doi.org/10.1109/ICSITech46713.2019.8987547>



- Ruuska, S., Hämmäläinen, W., Kajava, S., Mughal, M., Matilainen, P., & Mononen, J. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural Processes*, 148(January), 56–62. <https://doi.org/10.1016/j.beproc.2018.01.004>
- Sajid, N. A., Ahmad, M., Rahman, A. U., Zaman, G., Ahmed, M. S., Ibrahim, N., Ahmed, M. I. B., Krishnasamy, G., Alzaher, R., Alkharraa, M., AlKhulaifi, D., AlQahtani, M., Salam, A. A., Saraireh, L., Gollapalli, M., & Ahmed, R. (2023). A Novel Metadata Based Multi-Label Document Classification Technique. *Computer Systems Science and Engineering*, 46(2), 2195–2214. <https://doi.org/10.32604/csse.2023.033844>
- Salatino, A. A., Thanapalasingam, T., Mannocci, A., Osborne, F., & Motta, E. (2018). The computer science ontology: A large-scale taxonomy of research areas. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11137 LNCS*. Springer International Publishing. https://doi.org/10.1007/978-3-030-00668-6_12
- Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019). Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification. *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, 593–596. <https://doi.org/10.1109/ICACTM.2019.8776800>
- Singh, N. K., & Chand, S. (2022). Machine Learning-based Multilabel Toxic Comment Classification. *3rd IEEE 2022 International Conference on Computing, Communication, and Intelligent Systems, ICCIS 2022*, 435–439. <https://doi.org/10.1109/ICCIS56430.2022.10037626>
- Supriatman, A. (2021). Pembobotan TF-IDF pada Judul Penelitian Dosen Sebagai Dasar Klasifikasi Menggunakan Algoritma K-NN (Studi Kasus: Universitas Siliwangi). *Jurnal Serambi Engineering*, 6(1), 1573–1579. <https://doi.org/10.32672/jse.v6i1.2645>
- Suyanto. (2018). *Machine Learning Tingkat Dasar dan Lanjut*. Informatika Bandung.
- Trisari, W., Putri, H., Hendrowati, R., & Belakang, L. (2020). *Penggalan Teks Dengan Model Bag of Words Terhadap*. 2(1), 129–138.
- Ulumudin, G. I., Adiwijaya, A., & Mubarak, M. S. (2019). A multilabel classification on topics of qur’anic verses in English translation using K-Nearest Neighbor method with Weighted TF-IDF. *Journal of Physics: Conference Series*, 1192(1). <https://doi.org/10.1088/1742-6596/1192/1/012026>



- Wahyuni, R. T., Prastiyanto, D., & Suprpto, E. (2017). Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi. *Jurnal Teknik Elektro Universitas Negeri Semarang*, 9(1), 18–23.
- Wang, Z. W., Wang, S. K., Wan, B. T., & Song, W. W. (2020). A novel multi-label classification algorithm based on K-nearest neighbor and random walk. *International Journal of Distributed Sensor Networks*, 16(3). <https://doi.org/10.1177/1550147720911892>
- Wiraguna, A., Faraby, S. Al, & Adiwijaya. (2019). Klasifikasi Topik Multi Label pada Hadis Bukhari dalam Terjemahan Bahasa Indonesia Menggunakan Random Forest. *E-Proceeding of Engineering*, 6(1), 2144–2153.
- Xie, S., Hou, C., Yu, H., Zhang, Z., Luo, X., & Zhu, N. (2022). Multi-label disaster text classification via supervised contrastive learning for social media data. *Computers and Electrical Engineering*, 104(PA), 108401. <https://doi.org/10.1016/j.compeleceng.2022.108401>
- Yan, D., Li, K., Gu, S., & Yang, L. (2020). Network-Based Bag-of-Words Model for Text Classification. *IEEE Access*, 8, 82641–82652. <https://doi.org/10.1109/ACCESS.2020.2991074>
- Yang, Z., & Emmert-Streib, F. (2024). Optimal performance of Binary Relevance CNN in targeted multi-label text classification. *Knowledge-Based Systems*, 284(August 2023), 111286. <https://doi.org/10.1016/j.knosys.2023.111286>
- Yu, Y., Sun, Z., Sun, C., & Liu, W. (2021). Hierarchical Multilabel Text Classification via Multitask Learning. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, 2021-Novem*, 1138–1143. <https://doi.org/10.1109/ICTAI52525.2021.00180>
- Zhang, C., Mayr, P., Lu, W., & Zhang, Y. (2021). Knowledge Entity Extraction and Text Mining in the Era of Big Data. *Data and Information Management*, 5(3), 309–311. <https://doi.org/10.2478/dim-2021-0009>
- Zhang, M. L., & Zhou, Z. H. (2014). A review on multi-label learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 26(8), 1819–1837. <https://doi.org/10.1109/TKDE.2013.39>