DAFTAR PUSTAKA

Ahmed, N. A. (2024, November 10). What is A Confusion Matrix in Machine Learning? The Model Evaluation Tool Explained. *What is A Confusion Matrix in Machine Learning? The Model Evaluation Tool Explained*. https://www.datacamp.com/tutorial/what-is-a-confusion-matrix-in-machine-learning

Arta Dana, I. B. M., Hardjo, P. H., Marianti Purwanto, M. G., Pujiyanti, A. S., & Indriyani, I. (2021). Ethyl Methane Sulfonate (EMS) Effect on Mutagenesis in Balinese Red Rice (Oryza sativa cv. Barak Cenana). *Jurnal Biologi Tropis*, *21*(3), 698–705. https://doi.org/10.29303/jbt.v21i3.2815

Atagong, S. D., Tonnang, H., Senagi, K., Wamalwa, M., Agboka, K. M., & Odindi, J. (2025). A review on knowledge and information extraction from PDF documents and storage approaches. *Frontiers in Artificial Intelligence*, *8*, 1466092. https://doi.org/10.3389/frai.2025.1466092

Bokharaeian, B., Dehghani, M., & Diaz, A. (2023). Automatic extraction of ranked SNP-phenotype associations from text using a BERT-LSTM-based method. *BMC Bioinformatics*, *24*(1), 144. https://doi.org/10.1186/s12859-023-05236-w

Chen, L., Duan, L., Sun, M., Yang, Z., Li, H., Hu, K., Yang, H., & Liu, L. (2023). Current trends and insights on EMS mutagenesis application to studies on plant abiotic stress tolerance and development. *Frontiers in Plant Science*, *13*, 1052569. https://doi.org/10.3389/fpls.2022.1052569

Damayanti, F. (2021). Potensi Pemuliaan Mutasi Radiasi sebagai upaya Peningkatan Variasi Genetik pada Tanaman Hias. *EduBiologia: Biological Science and Education Journal*, *1*(2), 78. https://doi.org/10.30998/edubiologia.v1i2.9300

Deng, C., Zou, J., Deng, J., & Bai, M. (2021). Extraction of gene-disease association from literature using BioBERT. *The 2nd International Conference on Computing and Data Science*, 1–4. https://doi.org/10.1145/3448734.3450772.

Eang, C., & Lee, S. (2024). Improving the Accuracy and Effectiveness of Text Classification Based on the Integration of the Bert Model and a Recurrent Neural Network (RNN_Bert_Based). *Applied Sciences*, *14*(18), 8388. https://doi.org/10.3390/app14188388

EFSA Panel on Genetically Modified Organisms (GMO), Mullins, E., Bresson, J., Dalmay, T., Dewhurst, I. C., Epstein, M. M., Firbank, L. G., Guerche, P., Hejatko, J., Moreno, F. J., Naegeli, H., Nogué, F., Sánchez Serrano, J. J., Savoini, G., Veromann, E., Veronesi, F., Casacuberta, J., Lenzi, P., Munoz Guajardo, I., … Rostoks, N. (2021). In vivo and in vitro random mutagenesis techniques in plants. *EFSA Journal*, *19*(11). https://doi.org/10.2903/j.efsa.2021.6611

Faesol, N., Avivi, S., & Hartatik, S. (2022). Stabilitas Hasil Tiga Genotipe Tebu (Saccharum officinarum L.) Hasil Mutasi Ethyl Methane sulfonate (EMS). *Agrikultura*, *33*(3), 312. https://doi.org/10.24198/agrikultura.v33i3.40794

Farrell, M. J., Brierley, L., Willoughby, A., Yates, A., & Mideo, N. (2022). Past and future uses of text mining in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1975), 20212721. https://doi.org/10.1098/rspb.2021.2721

Firdaus, A., & Firdaus, W. I. (2021). *Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi: (Sebuah Ulasan)*. *13*(1).

Firmansyah, R. A. D. (2025). *DETEKSI KESESUAIAN PERLAKUAN PADA MUTAGENESIS EMS MENGGUNAKAN METODE RANDOM FOREST*.

González-Márquez, R., Schmidt, L., Schmidt, B. M., Berens, P., & Kobak, D. (2024). The landscape of biomedical research. *Patterns*, *5*(6), 100968. https://doi.org/10.1016/j.patter.2024.100968

Guo, Z., Wang, S., Wang, Y., Wang, Z., & Ou, G. (2024). A machine learning enhanced EMS mutagenesis probability map for efficient identification of causal mutations in Caenorhabditis elegans. *PLOS Genetics*, *20*(8), e1011377. https://doi.org/10.1371/journal.pgen.1011377

Irawan, G., Nilahayati, N., Nazimah, N., Handayani, R. S., & Nurdin, M. Y. (2023). Pengaruh Pemberian EMS (Ethyl Methane Sulfonate) Terhadap Pertumbuhan Kedelai (Glycine Max (L.) Merr.) Galur M.1.1.3. *Jurnal Ilmiah Mahasiswa Agroekoteknologi*, *1*(4), 87. https://doi.org/10.29103/jimatek.v1i4.10462

Kaliyar, R. K., Goswami, A., & Narang, P. (2021). FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimedia Tools and Applications*, *80*(8), 11765–11788. https://doi.org/10.1007/s11042-020-10183-2

Kasneci, E., Sessler, K., Fischer, F., Gasser, U., & Groh, G. (2023). *ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education*.

Keraghel, I., Morbieu, S., & Nadif, M. (2024). *Recent Advances in Named Entity Recognition: A Comprehensive Survey and Comparative Study* (arXiv:2401.10825). arXiv. https://doi.org/10.48550/arXiv.2401.10825

Kruiper, R., Vincent, J. F. V., Chen-Burger, J., Desmulliez, M. P. Y., & Konstas, I. (2020). *A Scientific Information Extraction Dataset for Nature Inspired Engineering*.

Laksono, F. P., & Fanata, W. I. D. (2022). PENGARUH INDUKSI MUTASI DENGAN MUTAGEN EMS (ETHYL METHANE SULFONATE) TERHADAP HASIL DAN KUALITAS KEDELAI HITAM (Glycine soja (L) Merrit). *Berkala Ilmiah Pertanian*, *5*(2), 120. https://doi.org/10.19184/bip.v5i2.29162

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, *36*(4), 1234–1240. https://doi.org/10.1093/bioinformatics/btz682

Lestari, E. G. (2021). *APLIKASI INDUKSI MUTASI UNTUK PEMULIAAN TANAMAN HIAS*.

Li, Y., Luan, Z., Liu, Y., Liu, H., Qi, J., & Han, D. (2024). Automated information extraction model enhancing traditional Chinese medicine RCT evidence extraction (Evi-BERT): Algorithm development and validation. *Frontiers in Artificial Intelligence*, *7*, 1454945. https://doi.org/10.3389/frai.2024.1454945

Lin, T., Wang, Y., Liu, X., & Qiu, X. (2022). A survey of transformers. *AI Open*, *3*, 111–132. https://doi.org/10.1016/j.aiopen.2022.10.001

Lo, K., Wang, L. L., Neumann, M., Kinney, R., & Weld, D. (2020). S2ORC: The Semantic Scholar Open Research Corpus. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4969–4983. https://doi.org/10.18653/v1/2020.acl-main.447

Mallick, R., Arnaboldi, V., Davis, P., Diamantakis, S., Zarowiecki, M., Howe, K., & Cb, C. (2022). Accelerated variant curation from scientific literature using biomedical text mining. *Open Access*.

Moeljani, I. R., Makziah, & Wahyuni, E. (2022). Estimation of Genetic Diversity and Ld50 Determination of Red Onion (Allium cepa var ascalonicum. Linn) Variety of Bauji

Results In Gamma Iradiation 60Co. *Jurnal Hortikultura Indonesia*, *12*(3), 183–190. https://doi.org/10.29244/jhi.12.3.183-190

Nemoto, S., Kitada, S., & Iyatomi, H. (2025). Majority or Minority: Data Imbalance Learning Method for Named Entity Recognition. *IEEE Access*, *13*, 9902–9909. https://doi.org/10.1109/ACCESS.2024.3522972

Nguyen, N. D., Tan, W., Du, L., Buntine, W., Beare, Ri., & Chen, C. (2023). AUC Maximization for Low-Resource Named Entity Recognition. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(11), 13389–13399. https://doi.org/10.1609/aaai.v37i11.26571

Nilahayati, N., Siregar, N. K., & Harianja, A. A. (2024). Mutagen Ethyl Methane Sulphonate Menginduksi Keragaman Genetik Kedelai (Glycine max (L.)Merr.) Kultivar Gepak Kuning Pada Generasi M2. *Jurnal Agrium*, *21*(3), 247. https://doi.org/10.29103/agrium.v21i3.18695

Ompo, F. A. A., & Pakereng, M. A. I. (2024). Penerapan Text Mining Untuk Advertising Pada Data Tweets Zalora Indonesia Dengan Menggunakan Metode K-Means Clustering. *Progresif: Jurnal Ilmiah Komputer*, *20*(1), 314. https://doi.org/10.35889/progresif.v20i1.1576

Park, Y.-J., Lee, M., Yang, G.-J., Park, S. J., & Sohn, C.-B. (2023). Web Interface of NER and RE with BERT for Biomedical Text Mining. *Applied Sciences*, *13*(8), 5163. https://doi.org/10.3390/app13085163

Putri, E. E., Pasangka, B., & Louk, A. C. (2024). *Pemuliaan Tanaman Bayam Lokal Timor dengan Metode Radiasi Multigamma Standar*. *4*(1).

Qasim, R., Bangyal, W. H., Alqarni, M. A., & Ali Almazroi, A. (2022). A Fine-Tuned BERT-Based Transfer Learning Approach for Text Classification. *Journal of Healthcare Engineering*, *2022*, 1–17. https://doi.org/10.1155/2022/3498123

Ramadhan, M. F., & Siswoyo, B. (2024). *Mengenal Model BERT dan Implementasinya untuk Analisis Sentimen Ulasan Game*.

Reddy, M. S., & Pinjari, O. B. (2023). Mutagenesis Approaches in Plants and their Applications. *Asian Journal of Agriculture and Allied Sciences*, *6*(1), 20–29. https://doi.org/10.56557/ajaas/2023/v6i135

Sani, A., Oktavio, A., Metasari, R., Santosa, T. A., Sjoraida, D. F., Rembe, E., Amri, M., & Guna, B. W. K. (2024). The Impact Machine Learning Algorithms: Study Meta-Analysis. *INTECOMS: Journal of Information Technology and Computer Science*, *7*(4), 1117–1123. https://doi.org/10.31539/intecoms.v7i4.10860

Shaveta. (2023). A review on machine learning. *International Journal of Science and Research Archive*, *9*(1), 281–285. https://doi.org/10.30574/ijsra.2023.9.1.0410

Simanullang, G. S. B., & The, J. A. (2024). Roles of Natural Language Processing in New Product Development Process: Literature Review. *Jurnal Rekayasa Sistem Industri*, *13*(1), 117–130. https://doi.org/10.26593/jrsi.v13i1.6790.117-130

Song, X., Salcianu, A., Song, Y., Dopson, D., & Zhou, D. (2021). Fast WordPiece Tokenization. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2089–2103. https://doi.org/10.18653/v1/2021.emnlp-main.160

Tinn, R., Cheng, H., Gu, Y., Usuyama, N., Liu, X., Naumann, T., Gao, J., & Poon, H. (2023). Fine-tuning large neural language models for biomedical natural language processing. *Patterns*, *4*(4), 100729. https://doi.org/10.1016/j.patter.2023.100729

Varela-Vega, A., Posada-Reyes, A.-B., & Méndez-Cruz, C.-F. (2024). Automatic extraction of transcriptional regulatory interactions of bacteria from biomedical literature using a BERT-based approach. *Database*, *2024*, baae094. https://doi.org/10.1093/database/baae094

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (arXiv:1706.03762). arXiv. https://doi.org/10.48550/arXiv.1706.03762

Wirahno, D. N., & Damayanti, W. (t.t.). *Pemanfaatan Text-Mining untuk Ekstraksi Informasi Artikel Ilmiah Bidang PAUD*.

Yang, J., Zhang, T., Tsai, C.-Y., Lu, Y., & Yao, L. (2024). Evolution and emerging trends of named entity recognition: Bibliometric analysis from 2000 to 2023. *Heliyon*, *10*(9), e30053. https://doi.org/10.1016/j.heliyon.2024.e30053

Yusoff, M. I. M. (2024). Machine Learning: An Overview. *Open Journal of Modelling and Simulation*, *12*(03), 89–99. https://doi.org/10.4236/ojmsi.2024.123006

Zhang, Z., & Chen, A. L. P. (2022). Biomedical named entity recognition with the combined feature attention and fully-shared multi-task learning. *BMC Bioinformatics*, *23*(1), 458. https://doi.org/10.1186/s12859-022-04994-3

Zhao, S., Su, C., Lu, Z., & Wang, F. (2021). Recent advances in biomedical literature mining. *Briefings in Bioinformatics*, *22*(3), bbaa057. https://doi.org/10.1093/bib/bbaa057

Zhao, X., Greenberg, J., An, Y., & Hu, X. T. (2021). Fine-Tuning BERT Model for Materials Named Entity Recognition. *2021 IEEE International Conference on Big Data (Big Data)*, 3717–3720. https://doi.org/10.1109/BigData52589.2021.9671697