

DAFTAR PUSTAKA

- Akca, O., Bayrak, G., Issifu, A. M., & Ganiz, M. C. (2022). Traditional Machine Learning and Deep Learning-based Text Classification for Turkish Law Documents using Transformers and Domain Adaptation. *16th International Conference on INnovations in Intelligent SysTems and Applications, INISTA 2022*. <https://doi.org/10.1109/INISTA55318.2022.9894051>
- Alexopoulos, C., Loukis, E., & Virkar, S. (2024). Architecture and Evaluation of an Advanced Legal Information Platform—Enhancing Productivity of Modern Legal Work. *Journal of the Knowledge Economy*, 15(2). <https://doi.org/10.1007/s13132-023-01415-5>
- Ali, A., Noah, S. A. M., & Zakaria, L. Q. (2023). *A BERT-Based model: Improving Crime News Documents Classification through Adopting Pre-trained Language Models*. <https://doi.org/10.21203/rs.3.rs-2582775/v1>
- AL-Qurishi, M. (2023). *Leveraging BERT Language Model for Arabic Long Document Classification*. <http://arxiv.org/abs/2305.03519>
- AL-Qurishi, M., AlQaseemi, S., & Soussi, R. (2022). AraLegal-BERT: A pretrained language model for Arabic Legal text. *NLLP 2022 - Natural Legal Language Processing Workshop 2022, Proceedings of the Workshop*. <https://doi.org/10.18653/v1/2022.nllp-1.31>
- Ansell, A., Ponti, E. M., Korhonen, A., & Vulić, I. (2023). Distilling Efficient Language-Specific Models for Cross-Lingual Transfer. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/10.18653/v1/2023.findings-acl.517>
- Atmani, H., & Moutacalli, M. T. (2024). Automating Document Classification in the Financial Markets: A Comparative Study of Simple Models Versus Complex Models. *2024 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET)*, 1–6.
- Aumiller, D., Almasian, S., Lackner, S., & Gertz, M. (2021). Structural text segmentation of legal documents. *Proceedings of the 18th International Conference on Artificial Intelligence and Law, ICAIL 2021*. <https://doi.org/10.1145/3462757.3466085>
- Bai, P., Xia, Y., & Xia, Y. (2020). Fusing Knowledge and Aspect Sentiment for Explainable Recommendation. *IEEE Access*, 8. <https://doi.org/10.1109/ACCESS.2020.3012347>
- Bambroo, P., & Awasthi, A. (2021). LegalDB: Long distilbert for legal document classification. *Proceedings of the 2021 1st International Conference on Advances in Electrical, Computing, Communications and Sustainable*

- Brasoveanu, A., Moodie, M., & Agrawal, R. (2020). Textual evidence for the perfunctoriness of independent medical reviews. *CEUR Workshop Proceedings*, 2657, 1–9. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>
- Carrasco Ramírez, J. G. (2024). Natural Language Processing Advancements: Breaking Barriers in Human-Computer Interaction. *Journal of Artificial Intelligence General science (JAIGS)* ISSN:3006-4023, 3(1). <https://doi.org/10.60087/jaigs.v3i1.63>
- Chaudhary, Y., Gupta, P., Saxena, K., Kulkarni, V., Runkler, T., & Schütze, H. (2020). TopicBERT for energy efficient document classification. *Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020*. <https://doi.org/10.18653/v1/2020.findings-emnlp.152>
- Clavié, B., & Alphonsus, M. (2021). The Unreasonable Effectiveness of the Baseline: Discussing SVMs in Legal Text Classification. *Frontiers in Artificial Intelligence and Applications*, 346. <https://doi.org/10.3233/FAIA210317>
- Dadalto, E., Romanelli, M., Pichler, G., & Piantanida, P. (2023). *A Data-Driven Measure of Relative Uncertainty for Misclassification Detection*. <http://arxiv.org/abs/2306.01710>
- Darji, H., Mitrović, J., & Granitzer, M. (2023). German BERT Model for Legal Named Entity Recognition. *International Conference on Agents and Artificial Intelligence*, 3. <https://doi.org/10.5220/0011749400003393>
- Ding, Z., Jiang, G., Zhang, S., Guo, L., & Lin, W. (2023). SKDBERT: Compressing BERT via Stochastic Knowledge Distillation. *Proceedings of the 37th AAAI Conference on Artificial Intelligence, AAAI 2023*, 37. <https://doi.org/10.1609/aaai.v37i6.25902>
- Dong, J. (2024). Natural Language Processing Pretraining Language Model for Computer Intelligent Recognition Technology. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 23(8). <https://doi.org/10.1145/3605210>
- Fang, Z., He, Y., & Procter, R. (2023). *CWTM: Leveraging Contextualized Word Embeddings from BERT for Neural Topic Modeling*. <http://arxiv.org/abs/2305.09329>
- Firma, Z., Pusat, A., Dan, P., Perkara, P., Konstitusi, M., Indonesia, R., Medan, J., Nomor, M. B., Pusat, J., Yulistiyaputri, R., & Penelitian, P. (2019). *ROMANTISME SISTEM HUKUM DI INDONESIA: KAJIAN ATAS KONSTRIBUSI HUKUM ADAT DAN HUKUM ISLAM TERHADAP*

PEMBANGUNAN HUKUM DI INDONESIA (The Romanticism of Legal Systems in Indonesia: The Study of The Contribution of Islamic Law And Islamic Law for Legal Development In Indonesia).

- Fragkogiannis, P., Forster, M., Lee, G. E., & Zhang, D. (2023). Context-Aware Classification of Legal Document Pages. *SIGIR 2023 - Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. <https://doi.org/10.1145/3539618.3591839>
- Gessler, L., & Zeldes, A. (2022). MicroBERT: Effective Training of Low-resource Monolingual BERTs through Parameter Reduction and Multitask Learning. *MRL 2022 - 2nd Workshop on Multi-Lingual Representation Learning, Proceedings of the Workshop*. <https://doi.org/10.18653/v1/2022.mrl-1.9>
- Haque, S., Eberhart, Z., Bansal, A., & McMillan, C. (2022). Semantic Similarity Metrics for Evaluating Source Code Summarization. *IEEE International Conference on Program Comprehension, 2022-March*, 36–47. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>
- Hegel, A., Shah, M., Peaslee, G., Roof, B., & Elwany, E. (2021). *The Law of Large Documents: Understanding the Structure of Legal Contracts Using Visual Cues*. <http://arxiv.org/abs/2107.08128>
- Islam, S., Elmekki, H., Elsebai, A., Bentahar, J., Drawel, N., Rjoub, G., & Pedrycz, W. (2024). A comprehensive survey on applications of transformers for deep learning tasks. Dalam *Expert Systems with Applications* (Vol. 241). <https://doi.org/10.1016/j.eswa.2023.122666>
- Jojoa, M., Eftekhar, P., Nowrouzi-Kia, B., & Garcia-Zapirain, B. (2024). Natural language processing analysis applied to COVID-19 open-text opinions using a distilBERT model for sentiment categorization. *AI and Society*, 39(3). <https://doi.org/10.1007/s00146-022-01594-w>
- Joloudari, J. H., Hussain, S., Nematollahi, M. A., Bagheri, R., Fazl, F., Alizadehsani, R., Lashgari, R., & Talukder, A. (2023). BERT-deep CNN: state of the art for sentiment analysis of COVID-19 tweets. *Social Network Analysis and Mining*, 13(1). <https://doi.org/10.1007/s13278-023-01102-y>
- Kamble, S., & Kasodekar, K. S. (2024). *Transformer Architecture for NetsDB*. <http://arxiv.org/abs/2405.04807>
- Kryeziu, L., & Shehu, V. (2023). Pre-Training MLM Using Bert for the Albanian Language. *SEEU Review*, 18(1). <https://doi.org/10.2478/seeur-2023-0035>
- Kumar, T., Mahrishi, M., & Sharma, G. (2023). Emotion recognition in Hindi text using multilingual BERT transformer. *Multimedia Tools and Applications*, 82(27). <https://doi.org/10.1007/s11042-023-15150-1>

- Kumari, V., Keshari, S., Sharma, Y., & Goel, L. (2022). Context-Based Question Answering System with Suggested Questions. *Proceedings of the Confluence 2022 - 12th International Conference on Cloud Computing, Data Science and Engineering*. <https://doi.org/10.1109/Confluence52989.2022.9734207>
- Lê, J. K., & Schmid, T. (2022). The Practice of Innovating Research Methods. *Organizational Research Methods*, 25(2). <https://doi.org/10.1177/1094428120935498>
- Li, B., & Wang, M. (2023). *Design of intelligent legal text analysis and information retrieval system based on BERT model*. <https://doi.org/10.21203/rs.3.rs-2994403/v1>
- Li, J., Zhang, L. L., Xu, J., Wang, Y., Yan, S., Xia, Y., Yang, Y., Cao, T., Sun, H., Deng, W., Zhang, Q., & Yang, M. (2023). Constraint-aware and Ranking-distilled Token Pruning for Efficient Transformer Inference. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. <https://doi.org/10.1145/3580305.3599284>
- Li, Y., Cao, J., Xu, Y., Zhu, L., & Dong, Z. Y. (2024). Deep learning based on Transformer architecture for power system short-term voltage stability assessment with class imbalance. *Renewable and Sustainable Energy Reviews*, 189. <https://doi.org/10.1016/j.rser.2023.113913>
- Liu, C., Zhu, W., Zhang, X., & Zhai, Q. (2023). Sentence part-enhanced BERT with respect to downstream tasks. *Complex and Intelligent Systems*, 9(1). <https://doi.org/10.1007/s40747-022-00819-1>
- Liu, Z., Xu, Z., Jin, J., Shen, Z., & Darrell, T. (2023). Dropout Reduces Underfitting. *Proceedings of Machine Learning Research*, 202.
- Lu, J., Henchion, M., Bacher, I., & Namee, B. Mac. (2021). A Sentence-Level Hierarchical BERT Model for Document Classification with Limited Labelled Data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12986 LNAI. https://doi.org/10.1007/978-3-030-88942-5_18
- Mahardika, A. G. (2023). Pengujian Satu Atap Sebagai Optimalisasi Penataan Regulasi dalam Sistem Ketatanegaraan Indonesia. *MORALITY: Jurnal Ilmu Hukum*, 9(1). <https://doi.org/10.52947/morality.v9i1.311>
- Maiya, A., Sricharan, I., Pandey, A., & S, S. K. (2021). *Tom: Leveraging trend of the observed gradients for faster convergence*. <http://arxiv.org/abs/2109.03820>
- Mamakas, D., Tsotsi, P., Androutsopoulos, I., & Chalkidis, I. (2022). Processing Long Legal Documents with Pre-trained Transformers: Modding LegalBERT and Longformer. *NLLP 2022 - Natural Legal Language Processing Workshop*

2022, *Proceedings of the Workshop*. <https://doi.org/10.18653/v1/2022.nllp-1.11>

Martinez, S. P., Zimmermann, M., Offermann, M. S., & Reither, F. (2024). *Exploring Genre and Success Classification through Song Lyrics using DistilBERT: A Fun NLP Venture*. <http://arxiv.org/abs/2407.21068>

Mielniczuk, J., & Wawrzęczyk, A. (2024). Augmented Prediction of a True Class for Positive Unlabeled Data Under Selection Bias. *Frontiers in Artificial Intelligence and Applications*, 392, 2725–2732. <https://doi.org/10.3233/FAIA240806>

Muffo, M., & Bertino, E. (2020). BERTino: An italian distilbert model. *CEUR Workshop Proceedings*, 2769. <https://doi.org/10.4000/books.aaccademia.8748>

Nair, T. (2023). Natural Language Processing used in sentiment analysis of poetry: a study of six common techniques. *Journal of Student Research*, 12(2). <https://doi.org/10.47611/jsrhs.v12i2.4418>

Nanni, L., Maguolo, G., & Lumini, A. (2021). Exploiting Adam-like Optimization Algorithms to Improve the Performance of Convolutional Neural Networks. *arXiv*.

Oh, J., Kim, M., Park, H., & Oh, H. (2023). Are You Depressed? Analyze User Utterances to Detect Depressive Emotions Using DistilBERT. *Applied Sciences (Switzerland)*, 13(10). <https://doi.org/10.3390/app13106223>

Owusu-Adjei, M., Hayfron-Acquah, J. Ben, Frimpong, T., & Abdul-Salaam, G. (2023). A systematic review of prediction accuracy as an evaluation measure for determining machine learning model performance in healthcare systems. *medRxiv*.

Pansambal, B. H., & Nandgaokar, A. B. (2023). Integrating Dropout Regularization Technique at Different Layers to Improve the Performance of Neural Networks. *International Journal of Advanced Computer Science and Applications*, 14(4). <https://doi.org/10.14569/IJACSA.2023.0140478>

Pardede, M. (2023). DILEMA REFORMASI REGULASI HUKUM INVESTASI DALAM SISTEM HUKUM SIPILDALAM PERSPEKTIF JAMINAN KEPASTIAN HUKUM PENANAMAN MODAL (Suatu Upaya Peningkatan Iklim Daya Saing dan Jaminan Kepastian Hukum Penanaman Modal). *Jurnal Penelitian Hukum De Jure*, 23(2). <https://doi.org/10.30641/dejure.2023.v23.231-244>

Park, J., Yi, D., & Ji, S. (2020). A novel learning rate schedule in optimization for neural networks and it's convergence. *Symmetry*, 12(4). <https://doi.org/10.3390/SYM12040660>

- Petrakova, M. S. (2022). Systematization of legislation as a way to ensure the effectiveness of legal regulation: theoretical and legal aspects. *Courier of Kutafin Moscow State Law University (MSAL)*, 1(11). <https://doi.org/10.17803/2311-5998.2021.87.11.233-240>
- Pradipta, D., & Widodo, E. (2024). *INFORMASI (Jurnal Informatika dan Sistem Informasi) Sentiment Analysis on Social Media using Bidirectional Encoder from Transformers (Case Study : Covid-19 Omicron)*.
- Quevedo, E., Rahman, M., Cerny, T., Rivas, P., & Bejarano, G. (2022). *Study of Question Answering on Legal Software Document using BERT based models*. <https://doi.org/10.52591/lxai202207103>
- Radiya-Dixit, E. (2020). How fine can fine-tuning be? Learning efficient language models. *Proceedings of Machine Learning Research*, 108.
- Riehl, K., Neunteufel, M., & Hemberg, M. (2023). Hierarchical confusion matrix for classification performance evaluation. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 72(5). <https://doi.org/10.1093/jrssc/qlad057>
- Saad Rasool, Israr Hanif, Qaisar Rasool, Humaira Afzal, & Mufti, M. R. (2025). A Text Mining Approach for Automated Case Classification of Judicial Judgment. *Proceedings of the Pakistan Academy of Sciences: A. Physical and Computational Sciences*, 62(1). [https://doi.org/10.53560/ppasa\(62-1\)880](https://doi.org/10.53560/ppasa(62-1)880)
- Salehin, I., & Kang, D. K. (2023). A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain. Dalam *Electronics (Switzerland)* (Vol. 12, Nomor 14). <https://doi.org/10.3390/electronics12143106>
- Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). *DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter*. <http://arxiv.org/abs/1910.01108>
- Shah, M. A., Iqbal, M. J., Noreen, N., & Ahmed, I. (2023). An Automated Text Document Classification Framework using BERT. *International Journal of Advanced Computer Science and Applications*, 14(3). <https://doi.org/10.14569/IJACSA.2023.0140332>
- Shaheen, Z., Wohlgenannt, G., & Filtz, E. (2020). *Large Scale Legal Text Classification Using Transformer Models*. <http://arxiv.org/abs/2010.12871>
- Shunk, J. (2022). *Neuron-Specific Dropout: A Deterministic Regularization Technique to Prevent Neural Networks from Overfitting & Reduce Dependence on Large Training Samples*. <http://arxiv.org/abs/2201.06938>

- Si, Y., & Roberts, K. (2021). *Three-level Hierarchical Transformer Networks for Long-sequence and Multiple Clinical Documents Classification*. <http://arxiv.org/abs/2104.08444>
- Sihotang, H. T., Albert, M., Riandari, F., & Rendell, L. (2023). Efficient optimization algorithms for various machine learning tasks, including classification, regression, and clustering. *Idea: Future Research*, 1(1). <https://doi.org/10.35335/idea.v1i1.3>
- Simanjuntak, E. (2019). Peran Yurisprudensi dalam Sistem Hukum di Indonesia. *Jurnal Konstitusi*, 16(1). <https://doi.org/10.31078/jk1615>
- Singh, H. V. P., & Mahmoud, Q. H. (2020). NLP-based approach for predicting HMI state sequences towards monitoring operator situational awareness. *Sensors (Switzerland)*, 20(11). <https://doi.org/10.3390/s20113228>
- Uchendu, A., & Le, T. (2024). *Unveiling Topological Structures in Text: A Comprehensive Survey of Topological Data Analysis Applications in NLP*. <http://arxiv.org/abs/2411.10298>
- Vatsal, S., Meyers, A., & Ortega, J. E. (2023). Classification of US Supreme Court Cases using BERT-Based Techniques. *International Conference Recent Advances in Natural Language Processing, RANLP*. https://doi.org/10.26615/978-954-452-092-2_128
- Wang, X., Weissweiler, L., Schütze, H., & Plank, B. (2023). How to Distill your BERT: An Empirical Study on the Impact of Weight Initialisation and Distillation Objectives. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2. <https://doi.org/10.18653/v1/2023.acl-short.157>
- Wang, Y., Kang, Y., Qin, C., Wang, H., Xu, Y., Zhang, Y., & Fu, Y. (2021). *Rethinking Adam: A Twofold Exponential Moving Average Approach*. <http://arxiv.org/abs/2106.11514>
- Wehnert, S., Dureja, S., Kutty, L., Sudhi, V., & De Luca, E. W. (2022). Applying BERT Embeddings to Predict Legal Textual Entailment. *The Review of Socionetwork Strategies*, 16(1). <https://doi.org/10.1007/s12626-022-00101-3>
- Xie, X., Xie, M., Moshayedi, A. J., & Noori Skandari, M. H. (2022). A Hybrid Improved Neural Networks Algorithm Based on L2 and Dropout Regularization. *Mathematical Problems in Engineering*, 2022. <https://doi.org/10.1155/2022/8220453>
- Zhang, R., Frei, S., & Bartlett, P. L. (2023). *Trained Transformers Learn Linear Models In-Context*. <http://arxiv.org/abs/2306.09927>

Zhang, Y., & Yang, Q. (2018). An overview of multi-task learning. Dalam *National Science Review* (Vol. 5, Nomor 1). <https://doi.org/10.1093/nsr/nwx105>

Zhao, W., Hu, H., Zhou, W., Shi, J., & Li, H. (2023). BEST: BERT Pre-training for Sign Language Recognition with Coupling Tokenization. *Proceedings of the 37th AAAI Conference on Artificial Intelligence, AAAI 2023*, 37. <https://doi.org/10.1609/aaai.v37i3.25470>