

DAFTAR PUSTAKA

- Alshamsan, A. R., & Chaudhry, S. A. (2022). Machine Learning Algorithms for Privacy Policy Classification: A Comparative Study. *2022 2nd IEEE International Conference on Software Engineering and Artificial Intelligence, SEAI 2022*, 214–219. <https://doi.org/10.1109/SEAI55746.2022.9832027>
- Alsuhailani, M. (2023). Deep learning-based sentence embeddings using Bert for textual entailment. *International Journal of Advanced Computer Science and Applications*, 14(8). <https://doi.org/10.14569/ijacsa.2023.01408108>
- Amaral, O., Abualhaija, S., Torre, D., Sabetzadeh, M., & Briand, L. C. (2021). *AI-enabled Automation for Completeness Checking of Privacy Policies*. <http://arxiv.org/abs/2106.05688>
- Amin, M. F. (2022). Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial. *Journal of Engineering Research*, 6(5), 0. <https://doi.org/10.21608/erjeng.2022.274526>
- Asif, M., Javed, Y., & Hussain, M. (2021). Automated Analysis of Pakistani Websites' Compliance with GDPR and Pakistan Data Protection Act. *Proceedings - 2021 International Conference on Frontiers of Information Technology, FIT 2021*, 234–239. <https://doi.org/10.1109/FIT53504.2021.00051>
- Breaux, T., & Antón, A. (2008). Analyzing Regulatory Rules for Privacy and Security Requirements. *IEEE Transactions on Software Engineering*, 34, 520. <https://doi.org/10.1109/TSE.2007.70746>.
- Cueto, M. del. (2025). *Grid search in python from scratch- hyperparameter tuning*. Towards Data Science. <https://towardsdatascience.com/grid-search-in-python-from-scratch-hyperparameter-tuning-3cca8443727b/>
- Del Alamo, J., Guamán, D., García, B., & Diez, A. (2022). A systematic mapping study on automated analysis of privacy policies. *Computing*, 104, 2053 - 2076. <https://doi.org/10.1007/s00607-022-01076-3>.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv [Cs.CL]*. Retrieved from <http://arxiv.org/abs/1810.04805>



- Fehér, T. (2020) *Fast Support Vector Classification with rapids cuml*, Medium. Available at: <https://medium.com/rapids-ai/fast-support-vector-classification-with-rapids-cuml-6e49f4a7d89e>.
- Gandhi, R. (2018). *Support Vector Machine - introduction to machine learning algorithms*. Medium. <https://medium.com/towards-data-science/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>
- Hamdani, R. el, Mustapha, M., Amariles, D. R., Troussel, A., Meeùs, S., & Krasnashchok, K. (2021). A combined rule-based and machine learning approach for automated GDPR compliance checking. *Proceedings of the 18th International Conference on Artificial Intelligence and Law, ICAIL 2021*, 40–49. <https://doi.org/10.1145/3462757.3466081>
- Hassani, S., Sabetzadeh, M., Amyot, D., & Liao, J. (2024). Rethinking Legal Compliance Automation: Opportunities with Large Language Models. *Proceedings of the IEEE International Conference on Requirements Engineering*, 432–440. <https://doi.org/10.1109/RE59067.2024.00051>
- Hosseini, M., Heaps, J., Slavin, R., Niu, J., & Breaux, T. (2021). Ambiguity and Generality in Natural Language Privacy Policies. 2021 IEEE 29th International Requirements Engineering Conference (RE), 70-81. <https://doi.org/10.1109/RE51729.2021.00014>
- Jain, A. (2024). *SVM Kernels and Its Type*. Medium. <https://medium.com/@abhishekjainindore24/svm-kernels-and-its-type-dfc3d5f2dcd8>
- Jain, N., & Kumar, R. (2022). A Review on Machine Learning & It's Algorithms. *International Journal of Soft Computing and Engineering*. <https://doi.org/10.35940/ijsc.e3583.1112522>.
- Juarta, B. (2023). Classification of Hoax News Using Machine Learning and Neural Networks with BERT Embeddings. *Proceedings - 2023 3rd International Conference on Electronic and Electrical Engineering and Intelligent System: Responsible Technology for Sustainable Humanity, ICE3IS 2023*, 311–316. <https://doi.org/10.1109/ICE3IS59323.2023.10335413>
- Kotal, A., Joshi, A., & Pande Joshi, K. (2021). The Effect of Text Ambiguity on creating Policy Knowledge Graphs. *2021 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing*,



Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), 1491–1500.

<https://doi.org/10.1109/ISPA-BDCloud-SocialCom-SustainCom52081.2021.00201>

Khanduja, D. K., & Kaur, S. (2023). The categorization of documents using support Vector Machines. *International Journal of Scientific Research in Computer Science and Engineering*, 11(6), 1–12.

<https://doi.org/10.26438/ijsrcse/v11i6.112>

Kumar, V. B., Ravichander, A., Story, P., & Sadeh, N. (2019). Quantifying the effect of in-domain distributed word representations: A study of privacy policies. *CEUR Workshop Proceedings*, 2335(2003), 46–52.

Khan, A. S., Khan, F. T., Mahmud, T., Khan, S. K., Sharmen, N., Hossain, M. S., & Andersson, K. (2024). Integrating BERT Embeddings with SVM for Prostate Cancer Prediction. *Proceedings - 6th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2024*, 574–579.

<https://doi.org/10.1109/ICEEICT62016.2024.10534547>

Li, X. H., & Liu, S. X. (2013). The Applications of Support Vector Machine in Natural Language Processing. *Applied Mechanics and Materials*, 2572–2575.
<https://doi.org/10.4028/WWW.SCIENTIFIC.NET/AMM.427-429.2572>

Liu, S., Zhao, B., Guo, R., Meng, G., Zhang, F., & Zhang, M. (2021). Have you been properly notified? automatic compliance analysis of privacy policy text with GDPR article 13. *The Web Conference 2021 - Proceedings of the World Wide Web Conference, WWW 2021*, 2154–2164.

<https://doi.org/10.1145/3442381.3450022>

Manandhar, S., Singh, K., & Nadkarni, A. (2024). Towards Automated Regulation Analysis for Effective Privacy Compliance. *Proceedings 2024 Network and Distributed System Security Symposium*.

<https://doi.org/10.14722/ndss.2024.24650>

Mousavi Nejad, N., Jabat, P., Nedelchev, R., Scerri, S., & Graux, D. (2020). Establishing a strong baseline for privacy policy classification. *IFIP Advances in Information and Communication Technology*, 370–383.
https://doi.org/10.1007/978-3-030-58201-2_25



- Mustapha, M., Krasnashchok, K., Al Bassit, A., & Skhiri, S. (2020). Privacy policy classification with XLNet (short paper). *Lecture Notes in Computer Science*, 250–257. https://doi.org/10.1007/978-3-030-66172-4_16
- Okumura, Y., Hirokawa, S., & Takeuchi, K. (2019). Significance of Low-Frequent Words in Concept Describing Document. *International Conference on Advanced Applied Informatics*, 1035–1036. <https://doi.org/10.1109/IIAI-AAI.2019.00214>
- Ravichandiran, S. (2021). Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT. Retrieved from <https://books.google.co.id/books?id=CvsWEAAAQBAJ>
- Reidenberg, J., Bhatia, J., Breaux, T., & Norton, T. (2016). Ambiguity in Privacy Policies and the Impact of Regulation. *The Journal of Legal Studies*, 45, S163 - S190. <https://doi.org/10.1086/688669>
- Salomão, A. (2023). *The Benefits of Pilot Testing in Research Studies*. Blog Mind the Graph. <https://mindthegraph.com/blog/pilot-testing-in-research/>
- Shalev-Shwartz, S. and Ben-David, S., 2014. *Understanding Machine Learning: From Theory to Algorithms*. New York: Cambridge University Press.
- Shen, J., Wang, Y., Wu, Q., & Zhao, Y. (2023). Privacy Policy Compliance Inspection of Financial Apps. *Proceedings - 2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security Companion, QRS-C 2023*, 216–219. <https://doi.org/10.1109/QRS-C60940.2023.00070>
- Silva, B., Denipitiyage, D., Seneviratne, S., Mahanti, A., & Seneviratne, A. (2024). Entailment-driven privacy policy classification with LLMS. *2024 Conference on Building a Secure & Empowered Cyberspace (BuildSEC)*, 8–15. <https://doi.org/10.1109/buildsec64048.2024.00010>
- Srinath, M., Wilson, S., & Giles, C. L. (2021). *Privacy at scale: Introducing the PrivaSeer corpus of web privacy policies*. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL)*. Association for Computational Linguistics.
- Usercentrics. (2023). *What is a privacy policy? documenting user privacy policies*. Consent Management Platform (CMP) Usercentrics. <https://usercentrics.com/knowledge-hub/what-is-a-privacy-policy-and-why-do-you-need-one/>



- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention Is All You Need. CoRR, abs/1706.03762. Retrieved from <http://arxiv.org/abs/1706.03762>
- Wright, G. (2023). *What is privacy policy?: Definition from TechTarget*. WhatIs. <https://www.techtarget.com/whatis/definition/privacy-policy>
- Wilson, S., Schaub, F., Dara, A. A., Liu, F., Cherivirala, S., Leon, P. G., Andersen, M. S., Zimmeck, S., Sathyendra, K. M., Russell, N. C., Norton, T. B., Hovy, E., Reidenberg, J., & Sadeh, N. (2016). *The Creation and Analysis of a Website Privacy Policy Corpus*. www.w3.org/2011/tracking-protection
- Xiang, A., Pei, W., & Yue, C. (2023). PolicyChecker: Analyzing the GDPR Completeness of Mobile Apps' Privacy Policies. *CCS 2023 - Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, 3373–3387. <https://doi.org/10.1145/3576915.3623067>
- Zhang, L., Moukafih, N., Alamri, H., Epiphaniou, G., & Maple, C. (2024). *A BERT-based Empirical Study of Privacy Policies' Compliance with GDPR*. <http://arxiv.org/abs/2407.06778>