UNIVERSITAS
GADJAH MADA
**Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan
Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM**
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR PUSTAKA

[1] S. Kemp, “Digital 2024: Indonesia,” DataReportal (in partnership with We Are Social and Meltwater), Feb 2024.

[2] V. Rideout dan M. B. Robb, “The Common Sense Census: Media Use by Kids Age Zero to Eight, 2020,” Common Sense Media, San Francisco, CA, 2020. Tersedia pada: https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-kids-age-zero-to-eight-2020

[3] N. Elias dan I. Sulkin, “YouTube Viewers in Diapers: An Exploration of Factors Associated with Amount of Toddlers’ Online Viewing,” *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, vol. 11, no. 3, 2017, doi: 10.5817/CP2017-3-2.

[4] J. S. Radesky, N. Kaciroti, H. M. Weeks, A. Schaller, dan A. L. Miller, “Longitudinal Associations Between Use of Mobile Devices for Calming and Emotional Reactivity and Executive Functioning in Children Aged 3 to 5 Years,” *JAMA Pediatrics*, vol. 177, no. 1, hlm. 62–70, 2023, doi: 10.1001/jamapediatrics.2022.4793.

[5] J. Piaget, *The Origins of Intelligence in Children*. New York: International Universities Press, 1952.

[6] A. Eshraghi, S. A. Moosavi, F. Haghighatdoost, A. Feizi, dan P. Adibi, “Is the Screen Time Duration Affecting Children’s Language Development? A Scoping Review,” *Clinical Epidemiology and Global Health*, vol. 22, hlm. 101216, 2023, doi: 10.1016/j.cegh.2023.101216.

[7] W. Jia, X. Deng, dan H. Zeng, “The Association between Screen Exposure and Social–Emotional Development in Children and Adolescents: A Meta-Analysis,” *Acta Psychologica*, vol. 261, hlm. 105818, Nov 2025, doi: 10.1016/j.actpsy.2025.105818.

[8] American Academy of Pediatrics, “Media and Young Minds,” *Pediatrics*, vol.

UNIVERSITAS
GADJAH MADA
**Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan
Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM**
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

138, no. 5, 2016, doi: 10.1542/peds.2016-2591.

[9] Republik Indonesia, "Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional." Lembaran Negara Republik Indonesia Tahun 2003 Nomor 78, 2003.

[10] World Health Organization, "Guidelines on Physical Activity, Sedentary Behaviour and Sleep for Children under 5 Years of Age," World Health Organization, Geneva, 2019. Tersedia pada: https://www.who.int/publications/i/item/9789241550536

[11] D. A. Christakis, F. J. Zimmerman, D. L. DiGiuseppe, dan C. A. McCarty, "Early Television Exposure and Subsequent Attentional Problems in Children," *Pediatrics*, vol. 113, no. 4, hlm. 708–713, 2004, doi: 10.1542/peds.113.4.708.

[12] K. Papadamou *dkk.*, "Disturbed YouTube for Kids: Characterizing and Detecting Inappropriate Videos Targeting Young Children," dalam *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, AAAI Press, 2020, hlm. 522–533. doi: 10.1609/icwsm.v14i1.7320.

[13] R. Tahir, F. Ahmed, H. Saeed, S. Ali, F. Zaffar, dan C. Wilson, "Bringing the Kid Back into YouTube Kids: Detecting Inappropriate Content on Video Streaming Platforms," dalam *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, ACM, 2019, hlm. 464–469. doi: 10.1145/3341161.3342913.

[14] K. Yousaf dan T. Nawaz, "A Deep Learning-Based Approach for Inappropriate Content Detection and Classification of YouTube Videos," *IEEE Access*, vol. 10, hlm. 16283–16298, 2022, doi: 10.1109/ACCESS.2022.3147519.

[15] B. M. Le *dkk.*, "Samba: Identifying Inappropriate Videos for Young Children on YouTube," dalam *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22)*, 2022, hlm. 88–97. doi: 10.1145/3511808.3557442.

UNIVERSITAS GADJAH MADA
Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

[16] V. Yekbote, J. U. Bhaskar, R. Yaswanthreddy, C. Vyshnavi, dan S. Shanti, "A Novel Approach for Inappropriate Content Detection and Classification of YouTube Videos Using Deep Learning," dalam *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 2024, hlm. 539–545. doi: 10.1109/ICAAIC60222.2024.10574927.

[17] M. Z. Ansari, F. Ahmad, F. Ahmad, S. M. Bilal, dan N. Fatima, "Offensive YouTube Comments Classification Using Deep Learning," dalam *2025 IEEE DELCON - International Conference on Recent Smart Technologies in Engineering for Sustainable Development*, 2025. doi: 10.1109/DELCON68055.2025.11400254.

[18] A. Vaswani *dkk.*, "Attention Is All You Need," dalam *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, 2017, hlm. 5998–6008.

[19] J. Devlin, M.-W. Chang, K. Lee, dan K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," dalam *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota: Association for Computational Linguistics, 2019, hlm. 4171–4186. doi: 10.18653/v1/N19-1423.

[20] V. Sanh, L. Debut, J. Chaumond, dan T. Wolf, "DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter." 2019. Tersedia pada: https://arxiv.org/abs/1910.01108

[21] A. Ezen-Can, "A Comparison of LSTM and BERT for Small Corpus." 2020. Tersedia pada: https://arxiv.org/abs/2009.05451

[22] S. M. Lundberg dan S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," dalam *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, 2017, hlm. 4765–4774.

[23] E. Strubell, A. Ganesh, dan A. McCallum, "Energy and Policy

UNIVERSITAS GADJAH MADA
Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Considerations for Deep Learning in NLP," dalam *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy: Association for Computational Linguistics, 2019, hlm. 3645–3650. doi: 10.18653/v1/P19-1355.

[24] S. Liang, "Comparative Analysis of SVM, XGBoost and Neural Network on Hate Speech Classification," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 5, hlm. 896–903, Okt 2021, doi: 10.29207/resti.v5i5.3506.

[25] F. A. Ramadhan, Rd. R. P. Ruslan, dan A. Zahra, "Sentiment Analysis of E-Commerce Product Reviews for Content Interaction Using Machine Learning," *Cakrawala – Repositori IMWI*, vol. 6, no. 1, hlm. 207–218, Feb 2023.

[26] A. Chaudhary dan S. Jain, "Enhancing Text Classification Using Hybrid Embeddings and Advanced Machine Learning Techniques," *International Journal of System Assurance Engineering and Management*, vol. 17, no. 1, hlm. 214–229, 2025, doi: 10.1007/s13198-025-02950-x.

[27] J.-W. Sun, J.-Q. Bao, dan L.-P. Bu, "Text Classification Algorithm Based on TF-IDF and BERT," dalam *2022 11th International Conference of Information and Communication Technology (ICTech)*, IEEE, 2022. doi: 10.1109/ICTech55460.2022.00112.

[28] Bisri dan R. Yunanda, "Sentiment Classification of GoFood Reviews with IndoBERT: A Case Study of Food Ordering Apps in Indonesia," dalam *2025 International Conference on Advanced Technologies in Energy and Informatic (ICATEI)*, 2025, hlm. 138–142. doi: 10.1109/ICATEI67676.2025.11405019.

[29] N. A. P. Masaling, Y. Arifin, H. Soeparno, dan F. Lumban Gaol, "Utilizing Sentiment Insights for Software Evaluation with Aspect-Based Sentiment Analysis in App Reviews Using IndoBERT-Lite," dalam *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 2024, hlm. 977–982. doi:

UNIVERSITAS
GADJAH MADA
**Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan
Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM**
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

10.1109/ISRITI64779.2024.10963515.

[30] S. Aftan dan H. Shah, "A Survey on BERT and Its Applications," dalam *2023 20th Learning and Technology Conference (L&T)*, IEEE, 2023. doi: 10.1109/LT58159.2023.10092289.

[31] T. Chen dan C. Guestrin, "XGBoost: A Scalable Tree Boosting System," dalam *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, hlm. 785–794. doi: 10.1145/2939672.2939785.

[32] C. Cortes dan V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, hlm. 273–297, 1995, doi: 10.1007/BF00994018.

[33] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," dalam *Machine Learning: ECML-98, 10th European Conference on Machine Learning*, dalam Lecture Notes in Computer Science, vol. 1398. Berlin, Heidelberg: Springer, 1998, hlm. 137–142. doi: 10.1007/BFb0026683.

[34] Md. M. Rahman, A. I. Shiplu, dan Y. Watanobe, "CommentClass: A Robust Ensemble Machine Learning Model for Comment Classification," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, hlm. 184, 2024, doi: 10.1007/s44196-024-00589-3.

[35] A. S. Xiao dan Q. Liang, "Spam Detection for YouTube Video Comments Using Machine Learning Approaches," *Machine Learning with Applications*, vol. 16, hlm. 100550, 2024, doi: 10.1016/j.mlwa.2024.100550.

[36] K. Vanguri dan S. Gorla, "YouTube Comments Spam Detection Model Comparison with Selective Feature Vectors and a Hybrid Metric," dalam *Proceedings of International Conference on Data Analytics and Insights (ICDAI 2024)*, dalam Lecture Notes in Networks and Systems, vol. 1233. Springer Nature Singapore, 2025, hlm. 559–576. doi: 10.1007/978-981-96-3287-9_44.

UNIVERSITAS
GADJAH MADA
**Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan
Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM**
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

[37] K. Srinivasan, V. Arul Mary Rexy, S. Choudhary, R. Melba Kani, P. Ravisankar, dan V. Chithra Devi, "Auto-Grading System for Higher Education Using BERT-XGBoost to Reduce Teacher Workload and Improve Academic Performance," dalam *2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS)*, 2025. doi: 10.1109/WorldSUAS66815.2025.11199294.

[38] Y. S. S. Rohit, T. Rushik, V. Hyndhava, dan P. B. Pati, "Classification in Viva Interactions: Distinguishing Questions, Answers, and Statements with BERT Embedding and SVM," dalam *2025 3rd IEEE International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, 2025. doi: 10.1109/ICKECS65700.2025.11034900.

[39] D. Jurafsky dan J. H. Martin, *Speech and Language Processing*, 3rd ed. Prentice Hall, 2023.

[40] G. Salton dan C. Buckley, "Term-Weighting Approaches in Automatic Text Retrieval," *Information Processing & Management*, vol. 24, no. 5, hlm. 513–523, 1988, doi: 10.1016/0306-4573(88)90021-0.

[41] F. Pedregosa *dkk.*, "Scikit-Learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, hlm. 2825–2830, 2011.

[42] M. Schuster dan K. Nakajima, "Japanese and Korean Voice Search," dalam *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2012, hlm. 5149–5152. doi: 10.1109/ICASSP.2012.6289079.

[43] B. Wilie *dkk.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," dalam *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing (AACL-IJCNLP 2020)*, Suzhou, China: Association for Computational Linguistics, Des 2020, hlm. 843–857. doi:

UNIVERSITAS
GADJAH MADA
**Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan
Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM**
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

10.18653/v1/2020.aacl-main.85.

[44] D. Sebastian, H. Purnomo, dan I. Sembiring, "BERT for Natural Language Processing in Bahasa Indonesia," dalam *2022 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, IEEE, 2022. doi: 10.1109/ICICyTA57421.2022.10038230.

[45] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, dan C.-J. Lin, "LIBLINEAR: A Library for Large Linear Classification," *Journal of Machine Learning Research*, vol. 9, hlm. 1871–1874, 2008.

[46] J. Bergstra dan Y. Bengio, "Random Search for Hyper-Parameter Optimization," *Journal of Machine Learning Research*, vol. 13, hlm. 281–305, 2012.

[47] S. M. Lundberg *dkk.*, "From Local Explanations to Global Understanding with Explainable AI for Trees," *Nature Machine Intelligence*, vol. 2, no. 1, hlm. 56–67, 2020, doi: 10.1038/s42256-019-0138-9.

[48] C. K. Chow, "On Optimum Recognition Error and Reject Tradeoff," *IEEE Transactions on Information Theory*, vol. 16, no. 1, hlm. 41–46, 1970, doi: 10.1109/TIT.1970.1054406.

[49] Y. Yao, "Three-Way Decisions with Probabilistic Rough Sets," *Information Sciences*, vol. 180, no. 3, hlm. 341–353, 2010, doi: 10.1016/j.ins.2009.09.021.

[50] R. T. Fielding, "Architectural Styles and the Design of Network-based Software Architectures," phdthesis, University of California, Irvine, 2000.

[51] S. Ramírez, "FastAPI: Modern, Fast (High-Performance) Web Framework for Building APIs with Python." 2026.

[52] Object Management Group (OMG), "Unified Modeling Language (UML) Specification, Version 2.5.1," Object Management Group, formal/17-12-05, Des 2017.

[53] M. Sokolova dan G. Lapalme, "A Systematic Analysis of Performance

UNIVERSITAS
GADJAH MADA
**Deteksi Konten YouTube Tidak Layak untuk Anak Batita (1 â€“ 3 Tahun) Berbahasa Indonesia Menggunakan
Ekstraksi Fitur TF-IDF dan IndoBERT dengan Classifier XGBoost dan SVM**
Alif Rizqullah Ma'ruf, Divi Galih Prasetyo Putri, S.Kom., M.Kom., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Measures for Classification Tasks," *Information Processing & Management*, vol. 45, no. 4, hlm. 427–437, 2009, doi: 10.1016/j.ipm.2009.03.002.

[54] H. K. N. Leung dan P. W. L. Wong, "A Study of User Acceptance Tests," *Software Quality Journal*, vol. 6, no. 2, hlm. 137–149, 1997, doi: 10.1023/A:1018503800709.

[55] R. Likert, "A Technique for the Measurement of Attitudes," *Archives of Psychology*, vol. 140, hlm. 1–55, 1932.

[56] Sudjana, *Metoda Statistika*, 6 ed. Bandung: Tarsito, 2005.

[57] Riduwan dan Akdon, *Rumus Dan Data Dalam Analisis Statistika*. Bandung: Alfabeta, 2007.

[58] J. Cohen, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, vol. 20, no. 1, hlm. 37–46, 1960, doi: 10.1177/001316446002000104.

[59] J. R. Landis dan G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, no. 1, hlm. 159–174, 1977, doi: 10.2307/2529310.

[60] J. Nielsen, *Usability Engineering*. San Francisco, CA: Morgan Kaufmann, 1994.