

BAB V

KESIMPULAN DAN SARAN

1.1 Kesimpulan

Berdasarkan analisis, perancangan, dan pengujian dari sistem pendeteksi situs yang mengandung konten negatif dengan menggunakan *Machine learning*, maka diperoleh beberapa kesimpulan:

1. Peneliti berhasil mengembangkan model klasifikasi konten negatif berbasis teks menggunakan algoritma *machine learning*. Model ini mampu mengidentifikasi dan mengelompokkan konten negatif website ke dalam kategori *gambling*, *pornografi*, *phishing*, dan konten positif yaitu *whitelist* secara *real-time*.
2. Model dengan akurasi tertinggi adalah *Support Vector Machine* (SVC) dengan rata-rata akurasi mencapai 82.24%. Hasil ini menunjukkan bahwa model SVM adalah pilihan terbaik untuk penelitian ini, mengungguli model-model lainnya dalam melakukan klasifikasi terhadap dataset yang digunakan.
3. Tingkat kesulitan dalam pembangunan sistem terletak pada proses web scraping untuk mengumpulkan dataset serta pada proses pelatihan model *Machine Learning* setelah fitur diekstraksi dari dataset tersebut.
4. Pengujian sistem menunjukkan tingkat keberhasilan sebesar 80% dalam mengklasifikasikan konten negatif. Meskipun belum sempurna, hasil ini menunjukkan bahwa model klasifikasi yang dikembangkan cukup efektif dalam meningkatkan keamanan jaringan dan mengurangi risiko paparan konten negatif.

1.2 Saran

Pada penelitian ini, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan sistem kedepannya:

1. Sebaiknya mengumpulkan lebih banyak data untuk dilatih dan diuji, terutama untuk kategori yang kurang terwakili atau memiliki tingkat kesalahan klasifikasi yang tinggi.
2. Untuk penelitian selanjutnya, dapat dipertimbangkan untuk menguji algoritma lain selain *Naive Bayes*, KNN, dan SVM, seperti *decision trees*, *random forests*, atau *deep learning*. Selain itu, teknik-teknik seperti *ensemble learning* atau *transfer learning* juga dapat dipertimbangkan untuk meningkatkan performa model lebih lanjut.
3. Sebaiknya model dikembangkan lebih lanjut dengan mengintegrasikannya ke dalam server untuk memungkinkan pengelolaan bandwidth dinamis berdasarkan riwayat akses konten negatif. Penelitian lebih lanjut diperlukan untuk mengembangkan sistem penalti bandwidth yang dinamis berdasarkan akses pengguna ke konten negatif, yang dapat dilacak melalui URL dan DNS.
4. Sebaiknya website perlu dikembangkan lebih lanjut dengan menambahkan sistem informasi yang berisikan infografis jumlah pengguna yang akses Wi-Fi yang terintergrasi kedalam server.
5. Untuk menghindari gangguan pada proses scraping yang disebabkan oleh Wi-Fi kampus yang memerlukan login ulang setiap 4 jam, sebaiknya gunakan koneksi internet dari kartu SIM. Hal ini akan membantu memastikan kelancaran dan kecepatan proses scraping yang seringkali memakan waktu cukup lama.
6. Untuk penelitian selanjutnya, melakukan labeling data yang tepat untuk memberikan data sesuai pada kategori agar memberikan hasil akurasi yang lebih tinggi.
7. Untuk meningkatkan kualitas pelabelan data dan mengurangi subjektivitas, libatkan 2 atau 3 orang dalam proses pelabelan. Kemudian, melakukan validasi silang untuk memastikan konsistensi dan akurasi label yang diberikan.

DAFTAR PUSTAKA

- [1] L. Yana Siregar, M. Irwan Padli Nasution Prodi Manajemen, and U. Negeri Islam Sumatera Utara, “Development of Information Technology on Increasing Business Online,” *J. Ilm. Manaj. dan Bisnis*, vol. 2, no. 1, pp. 71–75, 2020, [Online]. Available: <https://doi.org/10.30606/hjimbhttp://journal.upp.ac.id/index.php/Hirarki>
- [2] B. W. Aulia, M. Rizki, P. Prindiyana, and S. Surgana, “Peran Krusial Jaringan Komputer dan Basis Data dalam Era Digital,” *JUSTINFO | J. Sist. Inf. dan Teknol. Inf.*, vol. 1, no. 1, pp. 9–20, Dec. 2023, doi: 10.33197/justinfo.vol1.iss1.2023.1253.
- [3] R. R. Saputri and A. H. Yunial, “OKTAL : Jurnal Ilmu Komputer dan Science Implementasi Machine Learning Pada Sistem Pendeteksi URL Bermuatan Konten Negatif Menggunakan Metode Algoritma Naive Bayes Dan Support Vector Machine,” vol. 2, no. 11, pp. 3057–3062, 2023.
- [4] G. A. Sandag, J. Leopold, and V. F. Ong, “Klasifikasi Malicious Websites Menggunakan Algoritma K-NN Berdasarkan Application Layers dan Network Characteristics Malicious Websites Classification Using K-NN Algorithm Based on Application Layers and Network Characteristics,” vol. 4, no. 1, pp. 37–45, 2018.
- [5] A. Wibowo, “Sistem Deteksi Konten Negatif pada Teks Website Menggunakan Metode Random Forest,” *J. Informatics Data Sci.*, vol. 1, no. 2, pp. 2–6, 2023, doi: 10.24114/j-ids.v1i2.42737.
- [6] J. Janssen and W. Laatz, “Naive Bayes,” in *Statistische Datenanalyse mit SPSS*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 557–569. doi: 10.1007/978-3-662-53477-9_25.
- [7] Tatiane Machado, “No Title” *מיום והשקיי*, vol. 549, no. 1, pp. 40–42, 2017.
- [8] R. R. Saputri and A. H. Yunial, “OKTAL : Jurnal Ilmu Komputer dan Science Implementasi Machine Learning Pada Sistem Pendeteksi URL

- Bermuatan Konten Negatif Menggunakan Metode Algoritma Naive Bayes Dan Support Vector Machine,” vol. 2, no. 11, pp. 3057–3062, 2023, [Online]. Available: <https://journal.mediapublikasi.id/index.php/oktal>
- [9] K. R. Velpula, “Malicious URL Detection Using Machine Learning,” *Int. J. Food Nutr. Sci.*, vol. 11, no. 12, pp. 2063–2071, 2023, doi: 10.48047/ijfans/v11/i12/218.
- [10] B. Mahesh, “Machine Learning Algorithms - A Review,” *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/ART20203995.
- [11] N. Aini, R. Ruktiari, M. R. Pratama, and A. F. Buana, “Sistem Prediksi Tingkat Pencemaran Polusi Udara dengan Algoritma Naïve Bayes di Kota Makassar,” *Pros. Semin. Nas. Komun. dan Inform.*, vol. 3, pp. 83–90, 2019.
- [12] K. Kumar Mohbey and M. Acharya, “Python for Machine Learning,” *Basics Python Program. A Quick Guid. Beginners*, pp. 237–250, 2023, doi: 10.2174/9789815179637123010013.
- [13] D. Monett *et al.*, “Special Issue ‘On Defining Artificial Intelligence’—Commentaries and Author’s Response,” *J. Artif. Gen. Intell.*, vol. 11, no. 2, pp. 1–100, Feb. 2020, doi: 10.2478/jagi-2020-0003.
- [14] D. Azizah, A. Wibawa, and L. Budiarto, “Hakikat Epistemologi Artificial Intelligence,” *J. Inov. Teknol. dan Edukasi Tek.*, vol. 1, no. 8, pp. 592–598, Aug. 2021, doi: 10.17977/um068v1i82021p592-598.
- [15] A. Nassehi, R. Y. Zhong, X. Li, and B. I. Epureanu, “Review of machine learning technologies and artificial intelligence in modern manufacturing systems,” in *Design and Operation of Production Networks for Mass Personalization in the Era of Cloud Technology*, Elsevier, 2022, pp. 317–348. doi: 10.1016/B978-0-12-823657-4.00002-6.
- [16] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, “Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 153–160, Oct. 2023, doi:

10.57152/malcom.v3i2.897.

- [17] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning,” *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 1, Apr. 2021, doi: 10.28932/jutisi.v7i1.3216.
- [18] P. LANGLEY and S. SAGE, “Induction of Selective Bayesian Classifiers,” in *Uncertainty Proceedings 1994*, vol. 463, Elsevier, 1994, pp. 399–406. doi: 10.1016/B978-1-55860-332-5.50055-9.
- [19] J. Piironen, M. Paasiniemi, and A. Vehtari, “Projective inference in high-dimensional problems: Prediction and feature selection,” *Electron. J. Stat.*, vol. 14, no. 1, Jan. 2020, doi: 10.1214/20-EJS1711.
- [20] X.-Q. Liu, X.-C. Wang, L. Tao, F.-X. An, and G.-R. Jiang, “Alleviating conditional independence assumption of naive Bayes,” *Stat. Pap.*, vol. 65, no. 5, pp. 2835–2863, Jul. 2024, doi: 10.1007/s00362-023-01474-5.
- [21] A. J. S. Nurjanah and M. Akbar, “Sistem Pendukung Keputusan Penentuan Penerima Bantuan Siswa Miskin Metode Teorema Bayes,” *J. Inf. J. Penelit. dan Pengabd. Masy.*, vol. 6, no. 2, pp. 1–6, Dec. 2020, doi: 10.46808/informa.v6i2.176.
- [22] Z. Bi, Y. Han, C. Huang, and M. Wang, “Gaussian Naive Bayesian Data Classification Model Based on Clustering Algorithm,” in *Proceedings of the 2019 International Conference on Modeling, Analysis, Simulation Technologies and Applications (MASTA 2019)*, Paris, France: Atlantis Press, 2019. doi: 10.2991/masta-19.2019.67.
- [23] J. Laaksonen and E. Oja, “Classification with learning k-nearest neighbors,” in *Proceedings of International Conference on Neural Networks (ICNN’96)*, IEEE, pp. 1480–1483. doi: 10.1109/ICNN.1996.549118.
- [24] A. A. Thant and S. M. Aye, “Euclidean, Manhattan and Minkowski Distance Methods For Clustering Algorithms,” *Int. J. Sci. Res. Sci. Eng.*

- Technol.*, pp. 553–559, Jun. 2020, doi: 10.32628/IJSRSET2073118.
- [25] A. Putri *et al.*, “Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 1, pp. 20–26, May 2023, doi: 10.57152/malcom.v3i1.610.
- [26] S. Rizal, “Development of Big Data Analytics Model,” *ITEJ (Information Technol. Eng. Journals)*, vol. 4, no. 1, pp. 14–25, Jul. 2019, doi: 10.24235/itej.v4i1.47.
- [27] K. U. Syaliman, M. Zulfahmi, and A. A. Nababan, “Perbandingan Rapid Centroid Estimation (RCE) — K Nearest Neighbor (K-NN) Dengan K Means — K Nearest Neighbor (K-NN),” *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 1, pp. 79–89, Sep. 2017, doi: 10.30743/infotekjar.v2i1.166.
- [28] H. Christian Anderson Wint’s II, A. Irma Purnamasari, and T. Suprpti, “PREDIKSI HUNIAN HOTEL MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBORS,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1897–1903, Apr. 2024, doi: 10.36040/jati.v8i2.8342.
- [29] R. Safitri, I. Ali, and N. Rahaningsih, “ANALISIS SENTIMEN TERHADAP TREN FASHION DI MEDIA SOSIAL DENGAN METODE SUPPORT VECTOR MACHINE (SVM),” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1746–1754, Apr. 2024, doi: 10.36040/jati.v8i2.9045.
- [30] S. Saikin, S. Fadli, and M. Ashari, “Optimization of Support Vector Machine Method Using Feature Selection to Improve Classification Results,” *JISA(Jurnal Inform. dan Sains)*, vol. 4, no. 1, pp. 22–27, Jun. 2021, doi: 10.31326/jisa.v4i1.881.
- [31] A. R. Isnain, A. I. Sakti, D. Alita, and N. S. Marga, “SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM,” *J. Data Min. dan Sist. Inf.*, vol. 2, no. 1, p. 31, Feb. 2021, doi: 10.33365/jdmsi.v2i1.1021.

- [32] A. C. Muhammad *et al.*, “Dasar-dasar Pembelajaran Mesin,” p. 131, 2023.
- [33] P. Paisa, M. S. L, and A. Wahyuni, “PEMETAAN DAERAH RAWAN LONGSOR MENGGUNAKAN METODE PENCITRAAN SATELIT DI KABUPATEN ENREKANG SULAWESI SELATAN,” *J. GEOCELEBES*, pp. 80–90, Apr. 2021, doi: 10.20956/geocelebes.v5i1.12503.
- [34] D. Vikara, K. Bello, N. Wijaya, T. Warner, A. Sheriff, and D. Remson, “Evaluating the Impact of Proprietary Oil & Gas Data on Machine Learning Model Performance Using a Quasi-Experimental Analytical Approach,” Mar. 2022. doi: 10.2172/1855950.
- [35] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000. doi: 10.1017/CBO9780511801389.
- [36] M. A. Cano Lengua and E. A. Papa Quiroz, “A Systematic Literature Review on Support Vector Machines Applied to Classification,” in *2020 IEEE Engineering International Research Conference (EIRCON)*, IEEE, Oct. 2020, pp. 1–4. doi: 10.1109/EIRCON51178.2020.9254028.
- [37] H. Byun and S.-W. Lee, “Applications of Support Vector Machines for Pattern Recognition: A Survey,” 2002, pp. 213–236. doi: 10.1007/3-540-45665-1_17.
- [38] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, “An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review,” *Information*, vol. 15, no. 4, p. 235, Apr. 2024, doi: 10.3390/info15040235.
- [39] O. Okwuashi and C. E. Ndehedehe, “Deep support vector machine for hyperspectral image classification,” *Pattern Recognit.*, vol. 103, p. 107298, Jul. 2020, doi: 10.1016/j.patcog.2020.107298.
- [40] C. Savas and F. Dervis, “Comparative Performance Study of Linear and Gaussian Kernel SVM Implementations for Phase Scintillation Detection,” in *2019 International Conference on Localization and GNSS (ICL-GNSS)*,

- IEEE, Jun. 2019, pp. 1–6. doi: 10.1109/ICL-GNSS.2019.8752635.
- [41] P. Lubbers, B. Albers, and F. Salim, “Overview of HTML5,” in *Pro HTML5 Programming*, Berkeley, CA: Apress, 2011, pp. 1–22. doi: 10.1007/978-1-4302-3865-2_1.
- [42] M. A. Khder, “Web scraping or web crawling: State of art, techniques, approaches and application,” *Int. J. Adv. Soft Comput. its Appl.*, vol. 13, no. 3, pp. 144–168, 2021, doi: 10.15849/ijasca.211128.11.
- [43] S. Gupta, “Getting started with Web Scraping,” in *Web Scraping with Python*, Berkeley, CA: Apress, 2024. doi: 10.1007/979-8-8688-0776-3_1.
- [44] D. . Sarumathi, M. M. Sharmila, M. C.Saraswathy, and M. R. L. Priya, “The Benchmark Analysis of Different Web Scraping Tools and Techniques,” *IJARCCCE*, vol. 13, no. 5, May 2024, doi: 10.17148/IJARCCCE.2024.13596.
- [45] S. Rahman *et al.*, *Python : Dasar Dan Pemrograman Berorientasi Objek*. 2023.
- [46] S. Gowrishankar and A. Veena, “Parts of Python Programming Language,” in *Introduction to Python Programming*, Boca Raton : Taylor & Francis, a CRC title, part of the Taylor & Francis imprint, a member of the Taylor & Francis Group, the academic division of T&F Informa, plc, 2018.: Chapman and Hall/CRC, 2018, pp. 35–66. doi: 10.1201/9781351013239-2.
- [47] D. Sarkar, R. Bali, and T. Sharma, “The Python Machine Learning Ecosystem,” in *Practical Machine Learning with Python*, Berkeley, CA: Apress, 2018, pp. 67–118. doi: 10.1007/978-1-4842-3207-1_2.
- [48] I. Wahyudi and F. Alameka, “Analisis Blackbox Testing Dan User Acceptance Testing Terhadap Sistem Informasi Solusimedsosku,” *J. Teknosains Kodepena |*, vol. 04, no. 01, pp. 1–9, 2023.