

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Algoritma Naïve Bayes dan Support Vector Machine (SVM) mampu digunakan untuk mengklasifikasikan komentar spam promosi judi online pada platform YouTube. Berdasarkan hasil pengujian, kedua algoritma menghasilkan performa klasifikasi yang baik dengan nilai accuracy, precision, recall, dan F1-score yang tinggi. Namun, Support Vector Machine (SVM) menunjukkan performa yang lebih unggul dibandingkan Naïve Bayes pada sebagian besar metrik evaluasi.
2. Penerapan teknik data balancing menggunakan Random UnderSampling memberikan pengaruh terhadap kinerja klasifikasi komentar spam judi online. Hasil eksperimen menunjukkan bahwa data balancing mampu meningkatkan kemampuan model dalam mendeteksi kelas minoritas (Spam), yang ditunjukkan oleh peningkatan nilai recall. Namun, pada beberapa skenario terjadi penurunan nilai precision dan accuracy sehingga pengaruh data balancing perlu dipertimbangkan berdasarkan tujuan klasifikasi yang diinginkan.
3. Berdasarkan perbandingan seluruh skenario eksperimen, algoritma Support Vector Machine (SVM) tanpa penerapan data balancing menghasilkan performa terbaik dengan nilai Accuracy sebesar 99,54%, Precision sebesar 100%, Recall sebesar 96,54%, dan F1-score sebesar 98,23%. Dengan demikian, SVM tanpa data balancing merupakan model yang paling optimal untuk mendeteksi komentar spam promosi judi online pada dataset yang digunakan dalam penelitian ini.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya:

1. Penelitian selanjutnya dapat mengembangkan model klasifikasi dengan menggunakan algoritma *machine learning* lainnya, seperti Random Forest atau Logistic Regression, maupun pendekatan *deep learning* seperti Long Short-Term Memory (LSTM) atau Transformer, untuk membandingkan performa dengan metode yang digunakan dalam penelitian ini.
2. Dataset yang digunakan dapat diperluas baik dari segi jumlah maupun variasi sumber data, sehingga model yang dihasilkan memiliki kemampuan generalisasi yang lebih baik dan lebih robust dalam menghadapi berbagai pola komentar spam.
3. Penelitian selanjutnya dapat mengeksplorasi teknik penanganan *class imbalance* lainnya, seperti *oversampling* (SMOTE) atau kombinasi metode *hybrid sampling*, untuk menganalisis pengaruhnya terhadap performa model klasifikasi.
4. Pengembangan lebih lanjut juga dapat dilakukan pada tahap *preprocessing*, seperti penggunaan teknik *word embedding* atau *contextual embedding* (misalnya Word2Vec, FastText, atau BERT) untuk meningkatkan kualitas representasi teks dibandingkan metode TF-IDF.
5. Sistem yang dikembangkan dalam penelitian ini masih bersifat eksperimen berbasis dataset, sehingga pada penelitian selanjutnya dapat dikembangkan menjadi sistem deteksi spam secara real-time yang terintegrasi dengan platform media sosial.

REFERENSI

- [1] Simon Kemp, "Digital 2025: top social platforms in 2025," DataReportal. Accessed: Nov. 03, 2025. [Online]. Available: [https://datareportal.com/reports/digital-2025-sub-section-top-social-platforms?rq=social media](https://datareportal.com/reports/digital-2025-sub-section-top-social-platforms?rq=social%20media)
- [2] Kementerian Komunikasi dan Digital Republik Indonesia, "Perkuat Literasi Bahaya Judi Online, Komdigi Luncurkan Kampanye Keliling di 30 Kota." [Online]. Available: <https://www.komdigi.go.id/berita/siaran-pers/detail/perkuat-literasi-bahaya-judi-online-komdigi-luncurkan-kampanye-keliling-di-30-kota>
- [3] Kementerian Komunikasi dan Digital Republik Indonesia (corporate author), "Wamenkomdigi: Judi Online Musuh Besar Bangsa, 80 Ribu Anak Indonesia Terpapar." [Online]. Available: <https://www.komdigi.go.id/berita/berita-komdigi/detail/wamenkomdigi-judi-online-musuh-besar-bangsa-80-ribu-anak-indonesia-terpapar>
- [4] H. Attar, "Menko Polkam Ungkap 8,8 Juta Orang Main Judi Online: Ada Anak-anak hingga Anggota TNI-Polri." [Online]. Available: <https://nu.or.id/nasional/menko-polkam-ungkap-8-8-juta-orang-main-judi-online-ada-anak-anak-hingga-anggota-tni-polri-GKluK#:~:text=kondisi ini sudah dalam status,darurat>
- [5] Burhanudin, Y. Musa, and Y. Wihardi, "Klasifikasi Komentar Spam Pada Youtube Menggunakan Metode Naïve Bayes, Support Vector Machine, Dan K-Nearest Neighbors," *J. Inform. dan Komput.*, vol. 3, no. 2, pp. 54–59, 2108.

- [6] E. U. Ivan Rifky, *NATURAL LANGUAGE PROCESSING*. Yogyakarta: ANDI, 2023.
- [7] L. Efrizoni, S. Defit, M. Tajuddin, and A. Anggrawan, "Komparasi Ekstraksi Fitur dalam Klasifikasi Teks Multilabel Menggunakan Algoritma Machine Learning Comparison of Feature Extraction in Multilabel Text Classification Using Machine Learning Algorithm," vol. 21, no. 3, 2022, doi: 10.30812/matrik.v21i3.1851.
- [8] T. Widyanto, I. Ristiana, and A. Wibowo, "Komparasi Naïve Bayes dan SVM Analisis Sentimen RUU Kesehatan di Twitter," vol. 6, no. 3, pp. 147–161, 2023.
- [9] K. Lukman and S. Novianto, "Komparasi Algoritma Naïve Bayes dan SVM untuk Identifikasi Cyberbullying Selebriti di Media Sosial. Twitter," pp. 970–981, 2025, doi: 10.33364/algoritma/v.22-1.2196.
- [10] A. S. Xiao and Q. Liang, "Machine Learning with Applications Spam detection for Youtube video comments using machine learning approaches," vol. 16, no. April, 2024.
- [11] K. Kamdan, M. P. Anugrah, M. J. Almutaali, R. Ramdani, and I. L. Kharisma, "Performance Analysis of IndoBERT for Detection of Online Gambling Promotion in YouTube Comments †," 2025.
- [12] M. Clinton, T. Manullang, A. Z. Rakhman, H. Tantriawan, and A. Setiawan, "Comparative Analysis of CNN , Transformers , and Traditional ML for Classifying Online Gambling Spam Comments in Indonesian," vol. 9, no. 3, pp. 592–602, 2025.

- [13] B. A. Santoso, B. I. Nugroho, and D. U. Asyfiya, "TIN : Terapan Informatika Nusantara Perbandingan Algoritma Naïve Bayes , Support Vector Machine , dan Random Forest Untuk Analisis Sentimen Komentar Politik Youtube TIN : Terapan Informatika Nusantara," vol. 6, no. 4, pp. 381–390, 2025, doi: 10.47065/tin.v6i4.8326.
- [14] G. Airlangga, "Spam Detection on YouTube Comments Using Advanced Machine Learning Models : A Comparative Study," vol. 4, no. 2, pp. 500–508, 2024.
- [15] G. Airlangga, "Spam Detection in YouTube Comments Using Deep Learning Models : A Comparative Study of MLP , CNN , LSTM , BiLSTM , GRU , and Attention Mechanisms," vol. 4, no. October, pp. 1533–1538, 2024.
- [16] A. M. Isa, S. Ahmad, and N. M. Diah, "Detecting Offensive Malay Language Comments on YouTube using Support Vector Machine (SVM) and Naïve Bayes (NB) Model," vol. 6, no. 3, pp. 8548–8560, 2022.
- [17] G. Ignatow and R. Mihalcea, *An Introduction to Text Mining: Research Design, Data Collection, and Analysis*. SAGE Publications, 2017. [Online]. Available: <https://books.google.co.id/books?id=ubQkDwAAQBAJ>
- [18] T. Jo, *Text Mining: Concepts, Implementation, and Big Data Challenge*. in Studies in Big Data. Springer Nature Switzerland, 2024. [Online]. Available: <https://books.google.co.id/books?id=ewk4EQAAQBAJ>
- [19] D. Jurafsky and J. H. Martin, "Summary of Contents," 2025.

- [20] H. Lane and M. Dyshe, *Natural Language Processing in Action, Second Edition*. in In Action. Manning, 2025. [Online]. Available:
<https://books.google.co.id/books?id=WkIAEQAAQBAJ>
- [21] R. V Chandran, B. Venkateswaran, A. Kumar, and B. M. Krishna, *Foundations of Natural Language Processing*. Addition Publishing House, 2025. [Online]. Available:
<https://books.google.co.id/books?id=MnOTEQAAQBAJ>
- [22] C. C. Aggarwal, *Machine Learning for Text*. Cham, Switzerland: Springer International Publishing, 2018.
- [23] M. Awaludin, *Mengenal AI: Machine Learning, Deep Learning, dan Natural Language Processing*. Deepublish, 2025. [Online]. Available:
<https://books.google.co.id/books?id=5y51EQAAQBAJ>
- [24] G. Rebala, A. Ravi, and S. Churiwala, *An Introduction to Machine Learning*. Springer International Publishing, 2019. [Online]. Available:
<https://books.google.co.id/books?id=u8OWDwAAQBAJ>
- [25] Kusnawi, *Belajar Mudah dan Singkat MACHINE LEARNING*. Yogyakarta: ANDI, 2024.
- [26] D. Pandey, K. Niwaria, and B. Chourasia, "Machine Learning Algorithms : A Review," pp. 916–922, 2019.
- [27] M. Jordan, J. Kleinberg, and B. Scho, *PATTERN RECOGNITION AND MACHINE LEARNING*.
- [28] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode

Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter,” vol. 10, pp. 71–76, 2020.

- [29] Suyanto, *MACHINE LEARNING TINGKAT DASAR DAN LANJUT*, 2nd ed. Bandung: Informatika Bandung, 2022.

