

## DAFTAR ISI

<b>HALAMAN PENGESAHAN PEMBIMBING.....</b>	<b>II</b>
<b>HALAMAN PENGESAHAN PENGUJI.....</b>	<b>III</b>
<b>SURAT PERNYATAAN KARYA ASLI TUGAS AKHIR.....</b>	<b>IV</b>
<b>PERNYATAAN BEBAS PLAGIASI.....</b>	<b>V</b>
<b>KATA PENGANTAR.....</b>	<b>VI</b>
<b>ABSTRAK.....</b>	<b>VII</b>
<b>ABSTRACT.....</b>	<b>VIII</b>
<b>DAFTAR ISI.....</b>	<b>IX</b>
<b>DAFTAR GAMBAR.....</b>	<b>XI</b>
<b>DAFTAR MODUL PROGRAM.....</b>	<b>XII</b>
<b>DAFTAR TABEL.....</b>	<b>XIII</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Tugas Akhir.....	3
1.6. Tahapan Penelitian.....	3
1.6.1. Metodologi Penelitian.....	4
1.6.2. Metodologi Pengembangan Sistem.....	4
1.7. Sistematika Penulisan.....	5
<b>BAB II TINJAUAN LITERATUR.....</b>	<b>6</b>
2.1. Analisis Sentimen.....	6
2.2. X (Twitter).....	6
2.3. <i>Web Scrapping</i> .....	6
2.4. <i>Text Preprocessing</i> .....	6
2.4.1. <i>Cleaning</i> .....	6
2.4.2. <i>Case Folding</i> .....	6
2.4.3. <i>Tokenization</i> .....	6
2.4.4. <i>Word Normalization</i> .....	7
2.4.5. <i>Negation Handling</i> .....	7
2.4.6. <i>Stopword Removal</i> .....	7
2.4.7. <i>Stemming</i> .....	7
2.5. Pelabelan.....	7
2.6. TF-IDF.....	8
2.7. <i>Decision Tree</i> .....	8
2.8. <i>Algoritma Random Forest</i> .....	10
2.9. <i>Confusion Matrix</i> .....	11

2.10.	<i>State of The Art</i> .....	12
<b>BAB III METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM .....</b>		<b>16</b>
3.1.	Metodologi Penelitian .....	16
3.1.1.	Analisis Masalah .....	17
3.1.2.	Studi Literatur .....	17
3.1.3.	Pengumpulan Data .....	17
3.1.4.	Pelabelan Sentimen Manual .....	17
3.1.5.	<i>Text Preprocessing</i> .....	18
3.1.6.	Pelabelan Sentimen dengan <i>InSet Lexicon</i> .....	22
3.1.7.	Pembobotan TF-IDF .....	24
3.1.8.	Pembuatan Model <i>Random Forest</i> .....	28
3.1.9.	Pengujian Confusion Matrix .....	32
3.2.	Metodologi Pengembangan Sistem .....	33
3.2.1	Communication .....	33
3.2.2	Quick Plan and Modelling Quick Design .....	34
3.2.3	Construction of Prototype .....	38
3.2.4	Deployment Delivery and Feedback .....	38
<b>BAB IV PEMBAHASAN .....</b>		<b>39</b>
4.1.	Hasil Penelitian .....	39
4.1.1.	Implementasi Analisis Sentimen Olimpiade Paris 2024 .....	39
4.1.1.1.	Pengumpulan Data .....	39
4.1.1.2.	<i>Text Preprocessing</i> .....	39
4.1.1.3.	Pelabelan Sentimen dengan <i>Inset Lexicon</i> .....	43
4.1.1.4.	Pembobotan TF-IDF .....	44
4.1.1.5.	<i>Random Forest</i> .....	44
4.1.1.6.	Prediksi .....	45
4.1.2.	Implementasi Perangkat Lunak .....	45
4.2.	Pengujian .....	52
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>57</b>
<b>DAFTAR PUSTAKA .....</b>		<b>58</b>

## DAFTAR PUSTAKA

- Adrian, M. R., Putra, M. P., Rafialdy, M. H., & Rakhmawati, N. A. (2021). Perbandingan Metode Klasifikasi Random Forest dan SVM Pada Analisis Sentimen PSBB. *Jurnal Informatika Upgris*, 7(1), 36–40.
- Afdhal, I., Kurniawan, R., Iskandar, I., Salambue, R., Budianita, E., & Syafria, F. (2022). Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia. *Jurnal Nasional Komputasi Dan Teknologi Informasi*, 5(1), 122–130. <https://doi.org/https://doi.org/10.32672/jnkti.v5i1.4004>
- Anggina, S., Setiawan, N. Y., & Bachtiar, F. A. (2022). Analisis Ulasan Pelanggan Menggunakan Multinomial Naïve Bayes Classifier dengan Lexicon-Based dan TF-IDF Pada Formaggio Coffee and Resto. *Is The Best Accounting Information Systems and Information Technology Business Enterprise This Is Link for OJS Us*, 7(1), 76–90. <https://doi.org/10.34010/aisthebest.v7i1.7072>
- Arifin, N., Enri, U., & Sulistiyowati, N. (2021). Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*.
- Awangga, R. M., & Khonsa', N. H. (2022). Analisis Performa Algoritma Random Forest dan Naive Bayes Multinomial pada Dataset Ulasan Obat dan Ulasan Film. *InComTech : Jurnal Telekomunikasi Dan Komputer*, 12(1), 60–70. <https://doi.org/10.22441/incomtech.v12i1.14770>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/a:1010933404324>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees* (1st edition). <https://doi.org/10.1201/9781315139470>
- Diandra Audiansyah, D., Eka Ratnawati, D., & Trias Hanggara, B. (2022). Analisis Sentimen Aplikasi MyXL menggunakan Metode Support Vector Machine berdasarkan Ulasan Pengguna di Google Play Store. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(8), 3987–3994.
- Elhan, A., Kusuma, M., Hardhienata, D., Herdiyeni, Y., Wijaya, S. H., & Adisantoso, J. (2022). Analisis Sentimen Pengguna Twitter terhadap Vaksinasi COVID-19 di Indonesia menggunakan Algoritme Random Forest dan BERT Sentiment Analysis of Twitter Users on COVID-19 Vaccines in Indonesia using Random Forest and BERT Algorithms. *Jurnal Ilmu Komputer & Agri-Informatika*, 9(2), 199–211. <https://jurnal.ipb.ac.id/index.php/jika>
- Fitri, E., Yuliani, Y., Rosyida, S., & Gata, W. (2020). Analisis Sentimen Terhadap Aplikasi Ruanguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Transformatika*, 18(1), 71–80.
- Gifari, O. I., Adha, M., Rifky Hendrawan, I., Freddy, F., & Durrand, S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *JIFOTECH (Journal of Information Technology)*, 2(1), 36–40.

- Jia, K., Zhu, Y., Zhang, Y., Liu, F., & Qi, J. (2022). International public opinion analysis of four olympic games: From 2008 to 2022. *Journal of Safety Science and Resilience*, 3(3), 252–262. <https://doi.org/10.1016/j.jnlssr.2022.03.002>
- Josi, A., & Andretti Abdillah, L. (2014). Penerapan Teknik Web Scraping pada Mesin Pencari Artikel Ilmiah. *ArXiv (Cornell University)*, 159–164. <https://arxiv.org/pdf/1410.5777>
- Koto, F., & Rahmaningtyas, G. Y. (2017). InSet Lexicon: Evaluation of a Word List for Indonesian Sentiment Analysis in Microblogs. *Proceedings Of The 2017 International Conference On Asian Language Processing, IALP*, 391–394. <https://doi.org/10.1109/ialp.2017.8300625>
- Leonardo, R., Pratama, J., & Chrisnatalis. (2020). Perbandingan Metode Random Forest Dan Naïve Bayes Dalam Prediksi Keberhasilan Klien Telemarketing. *Jurnal Penelitian Teknik Informatika*, 3(2), 455–459.
- Mahendrajaya, R., Buntoro, G. A., & Setyawan, M. B. (2019). Analisis Sentimen Pengguna GoPay Menggunakan Metode Lexicon-Based dan Support Vector Machine. *KOMPUTEK: Jurnal Teknik Universitas Muhammadiyah Ponorogo*, 3(2), 52–63. <http://studentjournal.umpo.ac.id/index.php/komputek>
- Meisya, T., Aulia, P., Arifin, N., & Mayasari, R. (2001). Perbandingan Kernel Support Vector Machine (SVM) dalam Penerapan Analisis Sentimen Vaksinisasi COVID-19. *SINTECH (Science and Information Technology) Journal*, 4(2), 139–145. <https://doi.org/10.31598>
- Miftahusalam, A., Febby Nuraini, A., Khoirunisa, A. A., & Pratiwi, H. (2022). Perbandingan Algoritma Random Forest, Naïve Bayes, dan Support Vector Machine Pada Analisis Sentimen Twitter Mengenai Opini Masyarakat Terhadap Penghapusan Tenaga Honorer. *Seminar Nasional Official Statistics 2022*, 563–572.
- Mufidah, F. S., Winarno, S., Alzami, F., Udayanti, E. D., & Sani, R. R. (2022). Analisis Sentimen Masyarakat Terhadap Layanan Shopeefood Melalui Media Sosial Twitter Dengan Algoritma Naïve Bayes Classifier. *JOINS (Journal of Information System)*, 7(1), 14–25. <https://doi.org/10.33633/joins.v7i1.5883>
- Pak, A., & Paroubek, P. (2020). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *Language Resources and Evaluation*. <http://tumblr.com>
- Qaiser, S., & Ramsha, A. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *International Journal of Computer Applications*, 181(1), 25–29. <https://doi.org/10.5120/ijca2018917395>
- Rahman, M. R., Diansyah, A. F., & Hanafi, H. (2023). Sentiment Analysis on the Shopee Application on Playstore Using the Random Forest Classification Method. *Inform : Jurnal Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 9(1), 20–24. <https://doi.org/10.25139/inform.v9i1.5465>

- Ratino, R., Noor Hafidz, N., Anggraeni, S., & Gata, W. (2020). Sentimen Analisis Informasi Covid-19 menggunakan Support Vector Machine dan Naïve Bayes. *Jurnal Jupiter*, 12(2).
- Ratnawati, L., & Sulistyanningrum, D. R. (2019). Penerapan Random Forest untuk Mengukur Tingkat Keparahan Penyakit pada Daun Apel. *Jurnal Sins Dan Seni ITS*, 8(2). <https://doi.org/10.12962/j23373520.v8i2.48517>
- Sagita, A., Faqih, A., Dwilestari, G., Siswoyo, B., & Pratama, D. (2023). Penerapan Metode Random Forest dalam Menganalisis Sentimen Pengguna Aplikasi CapCut di Google Play Store. *Jurnal Mahasiswa Teknik Informatika*, 7(6), 3307–3313.
- Septiani, D., & Isabela, I. (2022). Analisis Term Frequency-Inverse Document Frequency (TF-IDF) dalam Temu Kembali Informasi pada Dokumen Teks. *SINTESIA: Jurnal Sistem Dan Teknologi Informasi Indonesia*, 1(2), 81–88.
- Sya'bani, M. M., & Umilasari, R. (2018). Penerapan Metode Cosine Similarity dan Pembobotan TF/IDF pada Sistem Klasifikasi Sinopsis Buku di Perpustakaan Kejaksaan Negeri Jember. *JUSTINDO (Jurnal Sistem & Teknologi Informasi Indonesia)*, 3(1), 31–42.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37(2), 267–307. [https://doi.org/10.1162/coli\\_a\\_00049](https://doi.org/10.1162/coli_a_00049)
- Tamrizal, A. M., & Yaqin, A. (2022). Perbandingan Algoritma Naïve Bayes, K-Nearest Neighbors dan Random Forest untuk Klasifikasi Sentimen Terhadap BPJS Kesehatan pada Media Twitter. *InComTech : Jurnal Telekomunikasi Dan Komputer*, 12(1), 01–10. <https://doi.org/10.22441/incomtech.v12i1.13642>
- Tangirala, S. (2020). Evaluating the Impact of GINI Index and Information Gain on Classification using Decision Tree Classifier Algorithm. *IJACSA) International Journal of Advanced Computer Science and Applications*, 11(2). <https://doi.org/10.14569/ijacsa.2020.0110277>
- Tarecha, R. I., Wahyudi, F., & Jannah, U. M. (2022). Penanganan Negasi dalam Analisa Sentimen Bahasa Indonesia. *JUSIFOR: Jurnal Sistem Informasi Dan Informatika*, 1(1), 51–58. <https://doi.org/10.33379/jusifor.v1i1.1276>
- Tri Romadloni, N., Santoso, I., & Budilaksono, S. (2019). Perbandingan Metode Naive Bayes, KNN, dan Decision Tree terhadap Analisis Sentimen Transportasi KRL Commuter Line. *Jurnal IKRA-ITH Informatika*, 3(2), 1–9.
- Widjiyati, N. (2021). Implementasi Algoritme Random Forest Pada Klasifikasi Dataset Credit Approval. *Jurnal Janitra Informatika Dan Sistem Informasi*, 1(1), 1–7. <https://doi.org/10.25008/janitra.v1i1.118>
- Zhao, X., & Nie, X. (2021). Splitting choice and computational complexity analysis of decision trees. *Entropy*, 23(10). <https://doi.org/10.3390/e23101241>

Zheng, D. (2024). Sentiment Analysis for Film Reviews Based on Random Forest. *Science and Technology of Engineering Chemistry and Environmental Protection*, 1(7).  
<https://doi.org/10.61173/5t8epb44>