

KLASIFIKASI OPINI DARING IMPLEMENTASI TEKNOLOGI VIDEO UNDERSTANDING UNTUK KESELAMATAN KERJA INDUSTRI 4.0

Galih Pramudia¹, Ade Bastian^{2*}, Dadan Zaliluddin³

^{1,2,3}Program Studi Informatika, Fakultas Teknik, Universitas Majalengka

Email: galihuniverse@gmail.com¹, adebastian@unma.ac.id², [dadanz@unma.ac.id](mailto:dadan@unma.ac.id)³

ABSTRACT

The advancement of Industry 4.0 technologies has accelerated the adoption of artificial intelligence, including video understanding, to improve efficiency and workplace safety in industrial environments. However, the implementation of this technology has generated diverse public opinions, particularly on social media platforms such as X. This study aims to identify public perceptions regarding the use of video understanding to ensure workplace safety through online opinion classification. Data were collected from X using a web scraping technique based on selected keywords. The dataset then underwent a series of preprocessing steps, including case folding, filtering, tokenization, normalization, stopword removal, and stemming. Feature representation was carried out using TF-IDF, followed by classification with the Multinomial Naive Bayes algorithm. The model was chosen due to its suitability in handling short and unstructured text, which aligns with the characteristics of tweets. Model performance was evaluated using accuracy, precision, recall, F1-score, and a confusion matrix. The results show that the majority of public opinions are neutral, with a smaller proportion of positive and negative sentiments. The model achieved a high performance with 98% accuracy and strong precision, recall, and F1-score across all sentiment classes. These findings provide insights into public acceptance of video understanding in workplace safety and can serve as valuable input for industry stakeholders in designing transparent and adaptive implementation strategies. Furthermore, this study highlights the contribution of data science in understanding societal perceptions of digital technologies in the modern era.

Keywords: Data Science, Naïve Bayes, Preprocessing, Video Understanding, X

Riwayat Artikel :

Tanggal diterima : 21-10-2025

Tanggal revisi : 04-11-2025

Tanggal terbit : 08-12-2025

DOI :

<https://doi.org/10.31949/j-ensitec.v12i01.16227>

1. PENDAHULUAN

Era Industri 4.0 ditandai dengan pemanfaatan teknologi cerdas dalam berbagai sektor, termasuk keselamatan kerja. Salah satu teknologi yang kini banyak diperbincangkan adalah video understanding, yakni kemampuan sistem untuk memahami dan menafsirkan konten

visual secara otomatis. Teknologi ini dipandang mampu membantu deteksi dini risiko kecelakaan dan memastikan kepatuhan terhadap standar keselamatan kerja secara real-time (Tang et al., 2025). Dengan kemampuan tersebut, video understanding menjadi solusi potensial dalam

This is an open access article under the CC BY-4.0 license.



mengurangi kecelakaan kerja yang masih menjadi persoalan serius di dunia industri.

Kecelakaan kerja tidak hanya berdampak pada produktivitas, tetapi juga pada aspek sosial, hukum, dan psikologis bagi para pekerja. Seiring meningkatnya kesadaran akan hak-hak pekerja dan kewajiban etis serta legal perusahaan dalam menyediakan lingkungan kerja yang aman, upaya memperkuat program keselamatan kerja menjadi semakin penting (Yogie Ernanda, 2023). Oleh karena itu, teknologi berbasis kecerdasan buatan seperti video understanding mulai mendapatkan perhatian sebagai pendekatan modern untuk mendukung lingkungan kerja yang aman.

Meski demikian, adopsi teknologi baru terutama yang berkaitan dengan pemantauan otomatis dan sistem pengawasan digital kerap memunculkan beragam respons dari masyarakat. Isu privasi, etika penggunaan, dan kepercayaan terhadap sistem berbasis kecerdasan buatan menjadi sumber perdebatan publik (Fahira Munawar et al., 2022). Media sosial, khususnya X (Twitter), berperan sebagai ruang terbuka tempat masyarakat menyuarakan pendapatnya secara spontan dan dalam skala besar (Gumilang & Riyandi, 2023). Dengan demikian, persepsi publik terhadap teknologi ini penting untuk dipahami karena dapat memengaruhi keberhasilan implementasinya di lapangan.

Untuk mengamati persepsi publik secara objektif, analisis sentimen menjadi pendekatan yang relevan. Analisis sentimen merupakan proses komputasi untuk mengevaluasi opini, emosi, dan sikap dalam teks (Aftab et al., 2023), baik secara umum maupun pada aspek tertentu melalui aspect-based sentiment analysis (Giffari, 2022). Dalam mengolah data opini dalam skala besar, diperlukan metode klasifikasi yang efisien. Multinomial Naive Bayes termasuk algoritma yang efektif untuk menganalisis teks pendek seperti tweet. Keunggulannya telah dibuktikan pada berbagai penelitian, mulai dari klasifikasi ulasan aplikasi Netflix (Khoirunnisaa et al., 2024), pendeteksian penyalahgunaan SMS (Amaia et al., 2023), hingga sentimen media sosial (Winahyu & Suharjo, 2021). Pendekatan ini bekerja melalui perhitungan probabilitas dalam memperkirakan kelas suatu data (Martantoh & Yanih, 2022), sehingga banyak digunakan dalam analisis opini publik (Prayoga Siswono et al., 2024; Aponno, 2022).

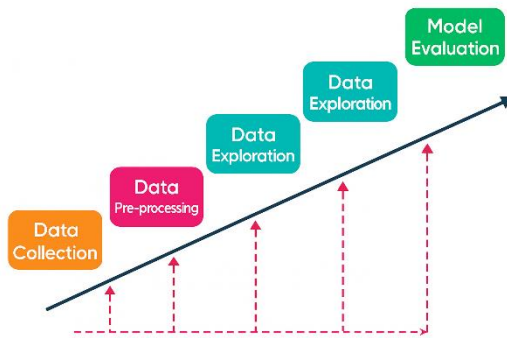
Sebelum proses klasifikasi dilakukan, data perlu melalui tahap preprocessing untuk meningkatkan kualitas teks agar siap diproses dalam model machine learning. Tahapan tersebut meliputi case folding untuk menyeragamkan huruf (Gustiara, 2023), filtering untuk menghapus angka, simbol, emotikon, dan tanda baca (Ramadhan & Mubarak, 2025), tokenization untuk memecah kalimat menjadi kata (Savitri et al., 2021), normalization untuk menyeragamkan kata tidak baku (Raif et al., 2024), stopword removal untuk menghilangkan kata umum yang tidak berkontribusi signifikan (Hendry et al., 2023), serta stemming untuk mengembalikan kata ke bentuk dasar (Soyusiawaty & Zahra, 2024). Representasi fitur dilakukan menggunakan TF-IDF yang dikenal efisien dan akurat dalam menentukan bobot kata penting (Putri Gabriella, 2023).

Dalam implementasinya, analisis dilakukan menggunakan Python karena memiliki pustaka lengkap dan mudah digunakan untuk pembangunan model machine learning (Gustiandi, 2023). Selain itu, Google Colab digunakan sebagai platform pemrograman berbasis komputasi awan yang mendukung pelatihan model secara daring (Vallejo et al., 2022). Dengan memanfaatkan media sosial sebagai sumber data opini publik dan algoritma pembelajaran mesin untuk klasifikasi, penelitian ini menganalisis persepsi masyarakat terkait penerapan teknologi video understanding dalam keselamatan kerja. Hasil penelitian diharapkan dapat memberikan gambaran objektif mengenai tingkat penerimaan publik serta menjadi masukan bagi industri dalam menyusun kebijakan implementasi yang lebih transparan dan adaptif (Yeung et al., 2021).

Penelitian ini bertujuan untuk mengklasifikasikan opini publik ke dalam kategori positif, netral, dan negatif menggunakan algoritma Multinomial Naive Bayes yang dikenal efektif untuk teks pendek seperti tweet (Aponno, 2022). Dengan memahami persepsi publik, perusahaan dapat menerapkan teknologi keamanan berbasis video secara lebih diterima dan bertanggung jawab oleh para pemangku kepentingan.

2. METODE PENELITIAN

Dalam penelitian ini, tahapan-tahapan dilaksanakan secara sistematis untuk memastikan bahwa setiap proses yang dilakukan dapat menghasilkan data dan analisis yang valid serta relevan terhadap tujuan penelitian. Tahapan tersebut dibagi menjadi 5 bagian utama yang mencakup proses pengumpulan data hingga evaluasi, yang tersaji pada gambar 1 dibawah ini :



Gambar 1. Tahapan Penelitian

Data penelitian dikumpulkan dari media sosial X menggunakan teknik web scraping dengan pustaka Python sncrape. Sebanyak 3.000 tweet berhasil diperoleh berdasarkan kata kunci “video understanding” dan “workplace safety”. Proses pembersihan dilakukan untuk memastikan kesesuaian dengan topik, sehingga tersisa 643 tweet yang dianggap relevan dengan penelitian ini.

Pra-pemrosesan teks dilakukan secara berurutan dengan beberapa tahap utama. Case folding digunakan untuk menyeragamkan huruf, filtering diterapkan untuk menghapus angka, tanda baca, emotikon, dan spasi ganda. Tahap tokenisasi memecah kalimat menjadi kata-kata tunggal, sementara normalisasi dilakukan untuk menyederhanakan bentuk kata yang tidak baku. Stopword removal menghapus kata umum yang tidak membawa makna signifikan, dan stemming menggunakan pustaka Sastrawi mengembalikan kata ke bentuk dasarnya. Setelah itu, representasi fitur dilakukan dengan metode TF-IDF agar setiap kata memiliki bobot berdasarkan frekuensi kemunculannya pada dokumen dan keseluruhan korpus.

Dataset yang telah dibersihkan dibagi menjadi dua bagian, yakni 514 data latih (80

persen) dan 129 data uji (20 persen). Model klasifikasi dibangun menggunakan algoritma Multinomial Naive Bayes yang cocok untuk dokumen teks pendek dengan distribusi kata bersifat diskret. Evaluasi dilakukan dengan mengukur akurasi, precision, recall, dan F1-score pada setiap kelas, serta menggunakan confusion matrix untuk menilai pola kesalahan klasifikasi.

3. HASIL DAN PEMBAHASAN

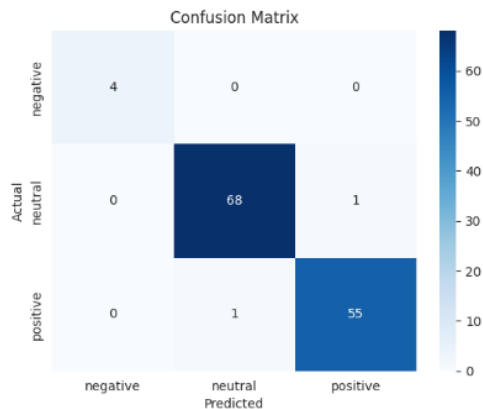
Pengolahan data dalam penelitian ini dilakukan untuk memperoleh informasi sentimen yang terkandung dalam tweet publik mengenai implementasi teknologi *video understanding* dalam konteks keselamatan kerja. Proses pengolahan mencakup tahapan transformasi data mentah menjadi bentuk yang dapat diinterpretasikan melalui teknik analisis teks dan klasifikasi berbasis machine learning. Berikut adalah contoh pengolahan data yang akan dilakukan misalkan terdapat tiga tweet berikut yang sudah dilabeli sentimennya secara manual seperti tersaji pada tabel 1 dibawah ini.

Tabel 1 Pelabelan Sentimen

Tweet	Teks Tweet	Label Sentimen
1	“video understanding membantu meningkatkan keselamatan kerja”	Positif
2	“pengawasan video berlebihan membuat pekerja tidak nyaman”	Negatif
3	“video digunakan di pabrik untuk memantau aktivitas kerja”	Netral

Setelah dilakukan proses seperti *case folding*, *filtering*, *tokenization*, dan *stopword removal*, seperti tersaji pada Tabel 2 dibawah ini.

Analisis confusion matrix memperlihatkan bahwa kesalahan klasifikasi umumnya terjadi pada sentimen netral yang terkadang dikategorikan sebagai positif atau negatif. Kesalahan ini dapat dipahami karena beberapa tweet informatif sering kali memiliki nuansa optimistis maupun kekhawatiran. Meski demikian, tingkat kesalahan sangat kecil, sehingga tidak memengaruhi performa keseluruhan model.



Gambar 6 Confusion Matrix

Berdasarkan hasil confusion matrix dari data uji 129, terdapat 4 negatif benar yang berhasil diprediksi oleh model, 68 netral berhasil diprediksi netral namun 1 netral diprediksi positif, dan 55 positif berhasil diprediksi positif tetapi 1 positif diprediksi netral. Jadi dari 129 ada 4 negatif, 69 netral, dan 56 positif. Temuan ini menunjukkan bahwa publik cenderung menerima teknologi video understanding sebagai sesuatu yang informatif dan bermanfaat, meskipun tetap ada kekhawatiran mengenai privasi. Bagi industri, hasil ini menjadi sinyal penting untuk menyusun kebijakan implementasi yang lebih inklusif, misalnya dengan memberikan jaminan perlindungan data dan penjelasan transparan kepada pekerja.

4. KESIMPULAN

Penelitian ini berhasil mengklasifikasikan opini publik terkait penerapan video understanding dalam keselamatan kerja menggunakan algoritma Multinomial Naive Bayes. Dari total 3.000 tweet, mayoritas opini bersifat netral, dengan dukungan positif yang cukup besar dan hanya sebagian kecil opini negatif. Setelah proses pembersihan, tersisa 643

data relevan yang dibagi menjadi 514 data latih dan 129 data uji. Model memberikan performa tinggi dengan akurasi 98 persen dan nilai F1 yang konsisten pada seluruh kelas. Temuan ini membuktikan efektivitas Naive Bayes dalam analisis sentimen teks pendek dan memberikan wawasan berharga bagi industri dalam mengelola strategi implementasi teknologi berbasis video understanding. Untuk penelitian selanjutnya, performa Multinomial Naive Bayes dapat dibandingkan dengan algoritma lain seperti SVM, Random Forest, atau model deep learning berbasis CNN, LSTM, dan BERT untuk mengetahui model yang paling optimal dalam analisis teks pendek. Peningkatan pra-pemrosesan, seperti stemming, lemmatization, TF-IDF, serta penanganan bahasa campuran dan ekspresi sarkastik juga berpotensi meningkatkan akurasi. Selain itu, hasil penelitian dapat diimplementasikan dalam bentuk dashboard analitik atau sistem pendukung keputusan yang mampu memantau sentimen secara real-time, sehingga lebih aplikatif bagi industri dalam mengevaluasi dan meningkatkan strategi penerapan teknologi video understanding pada keselamatan kerja.

5. REFERENSI

- [1] Ernanda, M. Y. (2023). Peningkatan Keamanan dan Keselamatan Kerja di Lingkungan Industri. *Circle Archive*, 1(3), 2–12. Link: <https://circle-archive.com/index.php/carc/article/view/58/49>
- [2] Tang, Y., Bi, J., Xu, S., Song, L., Liang, S., Wang, T., ... Xu, C. (2023). Video Understanding with Large Language Models: A Survey. *arXiv Preprint arXiv:2312.17432*. Link: <https://arxiv.org/abs/2312.17432>
- [3] Munawar, I. F., Alamanda, D. T. (2022). Sikap Masyarakat Terhadap Aplikasi PeduliLindungi Menggunakan Multiatribut Fishbein. *JPKP – Jurnal Pembangunan dan Kebijakan Publik*, 13(1), 1–19. Link: <https://jurnal.fisipuniga.ac.id/index.php/jp/article/view/132>
- [4] Khoirunnisaa, N., Nabila, K., Kesuma, N., Setiawan, S., Yunizar, A., & Yusuf, P.

- (2024). Klasifikasi Teks Ulasan Aplikasi Netflix pada Google Play Store Menggunakan Algoritma Naive Bayes dan SVM. *Jurnal Ilmu Komputer dan Teknologi Informasi*, 7(1), 64–73.
- [5] Amaia, E., Izzah, A. N., Akram, A., & Risal, N. (2023). Klasifikasi Penyalahgunaan Pesan Singkat Menggunakan Algoritma Naive Bayes. *Techno Xplore – Jurnal Ilmu Komputer dan Teknologi Informasi*.
- [6] Winahyu, J., & Suharjo, I. (2021). Aplikasi Web Analisis Sentimen dengan Algoritma Multinomial Naive Bayes. *KARMAPATI – Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika*, 10(2).
- [7] Aftab, F., Aftab, T., Khalid, H., et al. (2023). A Comprehensive Survey on Sentiment Analysis Techniques. *International Journal of Technology*, 14(6), 1288–1298. DOI: <https://doi.org/10.14716/ijtech.v14i6.6632>
- [8] Giffari, M. R. A. (2022). Analisis Sentimen Berbasis Aspek pada Ulasan Aplikasi Tangerang Live Menggunakan Latent Dirichlet Allocation dan Naive Bayes.
- [9] Siswono, A. P., et al. (2024). Analisis Sentimen Pelantikan Presiden Indonesia 2024 Menggunakan Model Klasifikasi dan Algoritma Naive Bayes. *Jurnal Sistem Informasi*, 4(1).
- [10] Gustiara, D. (2023). Implementasi Latent Dirichlet Allocation Terhadap Data Kasus Tindak Pidana.
- [11] Ramadhan, M. Z., & Mubarak, R. (2025). Analisis Sentimen Masyarakat Terhadap Kenaikan Bahan Bakar Minyak pada Media Sosial YouTube dengan Metode K-Nearest Neighbor dan Support Vector Machine.
- [12] Savitri, N. L. P. C., Rahman, R. A., Venyutzky, R., & Rakhmawati, N. A. (2021). Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning. *Jurnal Teknik Informatika dan Sistem Informasi*, 7(1). DOI: <https://doi.org/10.28932/jutisi.v7i1.3216>
- [13] Raif, M. I., Hidayati, N. N., & Matulatan, T. (2024). Otomatisasi Pendeteksi Kata Baku dan Tidak Baku pada Data Twitter Berbasis KBBI. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(2), 337–348. DOI: <https://doi.org/10.25126/jtiik.20241127404>
- [14] Hendry, M., Sianturi, F., Ridok, A., & Santoso, E. (2023). Peringkasan Teks Otomatis Menggunakan Metode Latent Semantic Analysis pada Artikel Berita Ekonomi Berbahasa Indonesia. *Jurnal Pengembangan Teknologi Informasi dan Komunikasi (J-PTIHK)*. Link: <https://j-ptiik.ub.ac.id/>
- [15] Soyusiawaty, D., & A.-Zahra. (2024). Peran Algoritma Stemming Nazief Adriani dalam Peningkatan Relevansi Pencarian Dokumen. *Jurnal Kesatria Informatika*, 5(1). DOI: <https://doi.org/10.30645/kesatria.v5i1.335>
- [16] Putri Gabriella, Y. A. (2023). Optimasi Penerimaan Siswa Baru dengan Penerapan Algoritma Text Mining dan TF-IDF. *Journal of Computing and Informatics Research*, 2(3), 110–117. DOI: <https://doi.org/10.47065/comforch.v2i3.941>
- [17] Gumilang, W., & Riyandi, A. (2023). Sentimen Analisis Pengguna Twitter terhadap SEA Games 2023 dengan Metode Naive Bayes. *Jurnal Akademika*, 16(1). DOI: <https://doi.org/10.53564/akademika.v16i1.1125>
- [18] Martantoh, E., & Yanih, N. (2022). Implementasi Metode Naive Bayes untuk Klasifikasi Karakteristik Kepribadian Siswa di Sekolah MTS Darussa'adah Menggunakan PHP MySQL.
- [19] Gustiandi, B. (2023). Langkah Awal Menguasai Bahasa Pemrograman Python. Penerbit BRIN. DOI: <https://doi.org/10.55981/brin.633>

- [20] Vallejo, W., Díaz-Urbe, C., & Fajardo, C. (2022). Google Colab and Virtual Simulations: Practical E-Learning Tools to Support the Teaching of Thermodynamics and to Introduce Coding to Students. *ACS Omega*, 7(8), 7421-7429. DOI: <https://doi.org/10.1021/acsomega.2c00362>
- [21] Yeung, A. W. K., et al. (2021). Implications of Twitter in Health-Related Research: A Landscape Analysis of the Scientific Literature. *Frontiers in Public Health*, 9. DOI: <https://doi.org/10.3389/fpubh.2021.654481>
- [22] Aponno, J. C. (2022). Penerapan Algoritma Sentiment Analysis dan Naïve Bayes terhadap Opini Pengunjung di Tempat Wisata Pantai Pintu Kota, Kota Ambon. *Jurnal Ilmiah*, 9(4), 3180–3188. Link: <https://jurnal.mdp.ac.id>