



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 18020-18028

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Sentimen Pengguna Media Sosial terhadap Penggunaan Artificial Intelligence dalam Dunia Kerja Menggunakan Naïve Bayes dan Support Vector Machine

Nur Azizah Harahap, Puspita Wanny, Khairunnisa

Magister Teknologi Infomasi, Pascasarjana, Universitas Panca Budi

nazizahhrp@gmail.com, puspitawanny142@gmail.com, nisak030720@gmail.com

Abstrak

Perkembangan Artificial Intelligence (AI) telah membawa perubahan signifikan dalam berbagai sektor pekerjaan dan memunculkan beragam tanggapan dari masyarakat. Sebagian masyarakat memandang AI sebagai peluang untuk meningkatkan produktivitas dan efisiensi kerja, sementara sebagian lainnya mengkhawatirkan dampaknya terhadap keberlangsungan lapangan pekerjaan. Penelitian ini bertujuan untuk menganalisis sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence dalam dunia kerja serta membandingkan kinerja algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam proses klasifikasi sentimen. Data penelitian diperoleh melalui proses scraping komentar YouTube yang membahas penggunaan AI dalam dunia kerja dan menghasilkan sebanyak 2.365 komentar. Data yang terkumpul melalui tahapan preprocessing yang meliputi case folding, cleansing, dan stopword removal, kemudian dilakukan pelabelan sentimen menggunakan pendekatan lexicon-based sentiment analysis. Selanjutnya, fitur diekstraksi menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) dan diklasifikasikan menggunakan algoritma Naïve Bayes dan Support Vector Machine. Hasil penelitian menunjukkan bahwa sentimen pengguna didominasi oleh sentimen netral sebanyak 68,37%, diikuti sentimen positif sebesar 20,04%, dan sentimen negatif sebesar 11,59%. Berdasarkan hasil pengujian, algoritma Naïve Bayes memperoleh accuracy sebesar 68,92%, precision sebesar 78,77%, recall sebesar 68,92%, dan F1-score sebesar 57,94%. Sementara itu, algoritma Support Vector Machine memperoleh accuracy sebesar 84,57%, precision sebesar 87,09%, recall sebesar 84,57%, dan F1-score sebesar 82,93%. Hasil tersebut menunjukkan bahwa Support Vector Machine memiliki performa yang lebih baik dibandingkan Naïve Bayes dalam melakukan klasifikasi sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence dalam dunia kerja.

Kata kunci: Analisis Sentimen, Artificial Intelligence, Naïve Bayes, Support Vector Machine, TF-IDF, Text Mining.

1. Pendahuluan

Perkembangan teknologi digital, khususnya Artificial Intelligence (AI), telah membawa perubahan besar dalam dunia kerja. AI mampu membantu manusia dalam menganalisis data, mengambil keputusan, dan meningkatkan efisiensi pekerjaan. Teknologi ini memberikan banyak manfaat melalui otomatisasi dan inovasi baru. Namun, perkembangan AI juga menimbulkan tantangan, seperti perubahan jenis pekerjaan, kebutuhan keterampilan baru, serta kekhawatiran beberapa pekerjaan manusia dapat tergantikan oleh sistem otomatis di masa depan [1].

Penerapan Artificial Intelligence (AI) dalam dunia kerja telah mengubah cara manusia berinteraksi dengan teknologi. Saat ini, banyak perusahaan menggunakan AI untuk membantu berbagai kegiatan, seperti layanan pelanggan melalui chatbot, analisis data, otomatisasi pekerjaan administrasi, hingga pembuatan teks, gambar, dan kode menggunakan teknologi generative AI [2]. Meskipun memberikan banyak manfaat dalam meningkatkan produktivitas, perkembangan AI juga menimbulkan kekhawatiran terkait perubahan lapangan pekerjaan dan kebutuhan keterampilan baru. Media sosial, seperti YouTube, menjadi tempat masyarakat menyampaikan pendapat mengenai dampak AI terhadap pekerjaan. Namun, banyaknya komentar yang tersedia membuat analisis secara manual menjadi sulit. Oleh karena itu, diperlukan metode otomatis untuk menganalisis opini masyarakat secara lebih cepat dan efektif.

Permasalahan tersebut menunjukkan bahwa dibutuhkan suatu metode otomatis yang mampu mengolah data teks dalam jumlah besar serta mengidentifikasi pola kecenderungan opini masyarakat secara lebih efektif. Salah satu

metode yang dapat diterapkan adalah analisis sentimen (sentiment analysis) [3]. Analisis sentimen merupakan bagian dari bidang Natural Language Processing (NLP) yang berfokus pada proses pengenalan, pengolahan, dan pengelompokan opini atau emosi yang terkandung dalam sebuah teks ke dalam beberapa kategori, seperti positif, negatif, maupun netral[4]. Melalui penerapan teknik text mining dan machine learning, analisis sentimen dapat digunakan untuk memahami persepsi masyarakat terhadap suatu isu tertentu secara lebih terstruktur, termasuk dalam mengkaji pandangan publik mengenai penerapan Artificial Intelligence (AI) dalam dunia kerja.

Dalam penelitian analisis sentimen, algoritma Naïve Bayes menjadi salah satu metode klasifikasi yang sering digunakan karena memiliki kemampuan dalam mengolah data teks secara efektif[5]. Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menerapkan Teorema Bayes untuk menghitung kemungkinan suatu data termasuk ke dalam kategori atau kelas tertentu berdasarkan pola data yang telah dipelajari[6]. Algoritma ini memiliki beberapa keunggulan, seperti proses komputasi yang relatif cepat, struktur metode yang sederhana, serta kemampuan yang baik dalam menangani data teks dengan jumlah fitur yang besar [7]. Salah satu pengembangan dari algoritma ini, yaitu Multinomial Naïve Bayes, banyak diterapkan dalam klasifikasi teks karena mampu bekerja dengan distribusi frekuensi kemunculan kata yang diperoleh melalui proses ekstraksi fitur, seperti Term Frequency-Inverse Document Frequency (TF-IDF)[8].

Selain Naïve Bayes, metode klasifikasi lain yang banyak digunakan dalam pengolahan data teks adalah Support Vector Machine (SVM). SVM merupakan algoritma supervised learning yang bekerja dengan menentukan hyperplane optimal untuk memisahkan data ke dalam beberapa kelas berdasarkan karakteristik fitur yang dimiliki[9]. Pendekatan berbasis machine learning seperti SVM juga banyak diterapkan dalam proses klasifikasi dan prediksi karena mampu menemukan pola pada data berdasarkan fitur yang telah diekstraksi[10]. Dalam analisis teks, SVM memiliki performa yang baik karena mampu menangani data dengan dimensi fitur yang tinggi, terutama setelah dilakukan proses representasi teks menggunakan metode TF-IDF. Penggunaan kernel linear pada SVM sering diterapkan dalam klasifikasi dokumen karena mampu menghasilkan batas pemisah yang efektif antara kelas sentimen yang berbeda[11].

Perbandingan antara algoritma Naïve Bayes dan Support Vector Machine menjadi topik yang menarik untuk dikaji karena kedua metode tersebut memiliki pendekatan dan karakteristik yang berbeda dalam melakukan klasifikasi. Naïve Bayes lebih unggul dalam hal efisiensi waktu dan kesederhanaan proses klasifikasi[12], sedangkan SVM memiliki kemampuan yang lebih baik dalam menemukan pola kompleks pada data berdimensi tinggi. Oleh karena itu, penelitian ini melakukan analisis perbandingan terhadap kedua algoritma tersebut untuk mengetahui metode yang memiliki tingkat performa terbaik dalam mengklasifikasikan sentimen pengguna media sosial mengenai penerapan Artificial Intelligence (AI) dalam dunia kerja.

Beberapa penelitian terdahulu telah menunjukkan bahwa algoritma Naïve Bayes dan SVM dapat diterapkan secara efektif dalam analisis sentimen berbasis media sosial. Penelitian yang dilakukan oleh Sari, Eviyanti, dan Astutik (2025) membahas penerapan analisis sentimen terhadap komentar pengguna YouTube menggunakan algoritma Naïve Bayes dan Support Vector Machine. Data penelitian diperoleh dari komentar pengguna YouTube yang kemudian melalui tahapan preprocessing seperti data cleaning, tokenisasi, normalisasi, dan penghapusan stopword. Selanjutnya data diklasifikasikan menggunakan metode machine learning untuk menentukan kategori sentimen pengguna. Hasil penelitian menunjukkan bahwa pendekatan machine learning mampu digunakan untuk mengidentifikasi pola opini pengguna pada data komentar YouTube yang memiliki karakteristik tidak terstruktur. Penelitian ini menunjukkan bahwa platform YouTube dapat menjadi sumber data yang efektif dalam memahami persepsi pengguna melalui pendekatan analisis sentimen[13].

Penelitian yang dilakukan oleh Kurniawan dan Widyassari (2025) dengan membahas perbandingan algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam melakukan klasifikasi sentimen terhadap komentar pengguna YouTube. Data penelitian diperoleh melalui proses crawling menggunakan YouTube Data API, kemudian dilakukan tahapan preprocessing yang meliputi case folding, cleaning, tokenizing, dan stopword removal. Setelah data diproses, ekstraksi fitur dilakukan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF), kemudian dilakukan klasifikasi menggunakan algoritma Naïve Bayes dan SVM. Evaluasi model dilakukan menggunakan metrik performa klasifikasi untuk mengetahui algoritma yang memiliki kemampuan terbaik dalam mengenali sentimen pengguna. Hasil penelitian menunjukkan bahwa penggunaan metode machine learning pada komentar YouTube dapat membantu mengidentifikasi opini publik terhadap suatu isu tertentu secara otomatis[14].

Selanjutnya, penelitian yang dilakukan oleh Permana et al. (2025) melakukan evaluasi terhadap performa kedua algoritma pada data komentar YouTube. Penelitian tersebut menerapkan tahapan preprocessing, ekstraksi fitur, dan proses klasifikasi untuk mengetahui kemampuan masing-masing metode. Hasil penelitian menunjukkan bahwa perbedaan karakteristik dataset dapat memengaruhi hasil klasifikasi, sehingga diperlukan pengukuran menggunakan beberapa metrik evaluasi, seperti accuracy, precision, recall, dan F1-score untuk memperoleh hasil analisis yang lebih objektif[15].

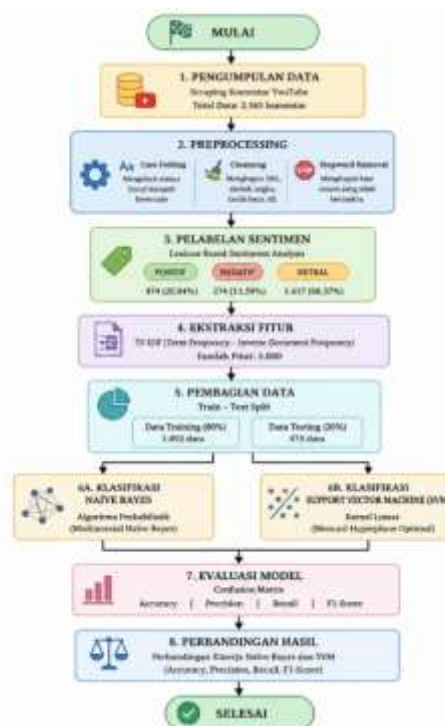
Berdasarkan beberapa penelitian terdahulu tersebut, dapat diketahui bahwa metode machine learning telah banyak digunakan dalam analisis sentimen untuk mengidentifikasi opini masyarakat pada berbagai platform media sosial. Namun, penelitian yang secara khusus membahas persepsi masyarakat terhadap penggunaan Artificial Intelligence dalam dunia kerja masih memiliki peluang untuk dikembangkan. Oleh karena itu, penelitian ini dilakukan untuk menganalisis sentimen komentar pengguna YouTube terkait penerapan AI dalam dunia kerja serta membandingkan performa algoritma Naïve Bayes dan Support Vector Machine. Hasil penelitian ini diharapkan dapat memberikan gambaran mengenai pandangan masyarakat terhadap perkembangan AI sekaligus menentukan algoritma yang lebih efektif dalam melakukan klasifikasi sentimen berbasis data teks.

2. Metode Penelitian

2.1. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan teknik text mining dan machine learning untuk menganalisis sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence (AI) dalam dunia kerja. Data penelitian diperoleh dari komentar pengguna YouTube yang membahas perkembangan AI dan dampaknya terhadap pekerjaan manusia. Analisis sentimen dilakukan untuk mengelompokkan komentar ke dalam kategori positif, negatif, dan netral, kemudian membandingkan performa algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam melakukan klasifikasi sentimen.

Tahapan penelitian dimulai dari pengumpulan data melalui proses scraping komentar YouTube, dilanjutkan dengan preprocessing data, pelabelan sentimen, ekstraksi fitur menggunakan TF-IDF, proses klasifikasi menggunakan algoritma Naïve Bayes dan SVM, serta evaluasi performa model menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Alur penelitian yang dilakukan ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

2.2. Pengumpulan dan Pengolahan Data

Data penelitian diperoleh melalui teknik scraping komentar pengguna YouTube pada beberapa video yang membahas penggunaan Artificial Intelligence dalam dunia kerja. Proses pengumpulan data dilakukan menggunakan bahasa pemrograman Python sehingga diperoleh sebanyak 2.365 komentar yang digunakan sebagai dataset penelitian.

Setelah data terkumpul, dilakukan proses preprocessing untuk membersihkan dan menyiapkan data sebelum dilakukan analisis. Tahapan preprocessing meliputi case folding, cleansing, dan stopword removal. Case folding dilakukan dengan mengubah seluruh huruf menjadi huruf kecil agar tidak terjadi perbedaan representasi kata akibat penggunaan huruf kapital. Cleansing dilakukan untuk menghapus karakter yang tidak diperlukan seperti URL, simbol, angka, dan tanda baca. Selanjutnya stopword removal digunakan untuk menghilangkan kata-kata umum yang tidak memiliki kontribusi signifikan terhadap pembentukan sentimen.

Data yang telah melalui preprocessing kemudian diberikan label sentimen menggunakan pendekatan lexicon-based sentiment analysis. Pelabelan dilakukan dengan memanfaatkan kamus kata positif dan negatif yang digunakan untuk menentukan kecenderungan sentimen suatu komentar. Hasil pelabelan menghasilkan tiga kategori sentimen, yaitu positif, negatif, dan netral. Sentimen positif merepresentasikan komentar yang mendukung atau memandang AI sebagai peluang, sentimen negatif merepresentasikan komentar yang menunjukkan kekhawatiran terhadap dampak AI, sedangkan sentimen netral menggambarkan komentar yang bersifat informatif atau tidak menunjukkan kecenderungan sentimen tertentu.

Tahap berikutnya adalah ekstraksi fitur menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk mengubah data teks menjadi representasi numerik sehingga dapat diproses oleh algoritma machine learning. TF-IDF memberikan bobot yang lebih tinggi pada kata-kata yang sering muncul dalam suatu dokumen tetapi jarang muncul pada keseluruhan dataset.

Secara matematis, nilai TF-IDF dihitung menggunakan persamaan:

Persamaan TF-IDF ditunjukkan pada Persamaan (3.1).

$$TFIDF(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

dengan nilai IDF dihitung menggunakan persamaan:

$$IDF(t) = \log \left(\frac{N}{df(t)} \right) \quad (2)$$

dimana N merupakan jumlah dokumen dan $df(t)$ merupakan jumlah dokumen yang mengandung kata t.

2.3. Klasifikasi dan Evaluasi Model

Dataset yang telah diekstraksi menggunakan TF-IDF kemudian dibagi menjadi data training dan data testing dengan perbandingan 80% dan 20%. Dari total 2.365 data, sebanyak 1.892 data digunakan sebagai data training dan 473 data digunakan sebagai data testing. Pembagian data dilakukan menggunakan metode train-test split untuk memastikan model dapat dievaluasi menggunakan data yang belum pernah digunakan pada proses pelatihan.

Proses klasifikasi dilakukan menggunakan dua algoritma, yaitu Naïve Bayes dan Support Vector Machine. Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen. Pada penelitian ini digunakan Multinomial Naïve Bayes karena sesuai untuk data teks hasil ekstraksi TF-IDF.

Persamaan Naïve Bayes ditunjukkan pada Persamaan (3.3).

$$P(C | X) = \frac{P(X|C) \times P(C)}{P(X)} \quad (3)$$

dimana $P(C|X)$ merupakan probabilitas suatu kelas berdasarkan fitur yang diamati.

Selain Naïve Bayes, penelitian ini juga menggunakan Support Vector Machine yang merupakan algoritma supervised learning untuk klasifikasi data. SVM bekerja dengan membangun hyperplane optimal yang memisahkan setiap kelas sehingga mampu menghasilkan kemampuan klasifikasi yang baik pada data berdimensi tinggi seperti data teks. Pada penelitian ini digunakan kernel linear karena lebih sesuai untuk proses klasifikasi dokumen berbasis TF-IDF.

Kinerja kedua algoritma dievaluasi menggunakan confusion matrix. Berdasarkan confusion matrix dihitung nilai accuracy, precision, recall, dan F1-score. Accuracy digunakan untuk mengukur tingkat ketepatan model secara keseluruhan, precision digunakan untuk mengukur ketepatan prediksi pada suatu kelas, recall digunakan untuk mengukur kemampuan model dalam menemukan seluruh data pada kelas tertentu, sedangkan F1-score merupakan rata-rata harmonis antara precision dan recall. Hasil evaluasi tersebut digunakan untuk menentukan algoritma yang memiliki performa terbaik dalam melakukan analisis sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence dalam dunia kerja.

3. Hasil dan Pembahasan

3.1. Hasil Penelitian

Pada penelitian ini data diperoleh melalui proses scraping komentar pengguna YouTube pada beberapa video yang membahas penggunaan Artificial Intelligence (AI) dalam dunia kerja. Proses pengumpulan data menghasilkan sebanyak 2.365 komentar yang selanjutnya digunakan sebagai dataset penelitian. Setelah data terkumpul, dilakukan tahapan preprocessing yang meliputi case folding, cleansing, dan stopword removal untuk menghasilkan data teks yang lebih bersih dan siap dianalisis.

Tahap berikutnya adalah pelabelan sentimen menggunakan pendekatan lexicon-based sentiment analysis. Hasil pelabelan menunjukkan bahwa mayoritas komentar berada pada kategori netral sebanyak 1.617 komentar atau 68,37% dari keseluruhan data. Sentimen positif berjumlah 474 komentar atau 20,04%, sedangkan sentimen negatif sebanyak 274 komentar atau 11,59%.

Tabel 1. Distribusi Sentimen

Sentimen	Jumlah	Persentase
Netral	1.617	68,37%
Positif	474	20,04%
Negatif	274	11,59%
Total	2.365	100%

Berdasarkan distribusi tersebut dapat diketahui bahwa sebagian besar pengguna media sosial memberikan tanggapan yang bersifat netral. Kondisi ini menunjukkan bahwa masyarakat masih berada pada tahap mengamati dan mendiskusikan perkembangan AI tanpa secara jelas menunjukkan dukungan ataupun penolakan terhadap penggunaannya dalam dunia kerja.

Setelah proses pelabelan selesai dilakukan, data ditransformasikan ke dalam bentuk numerik menggunakan metode TF-IDF. Hasil ekstraksi fitur menghasilkan 5.000 fitur yang digunakan sebagai masukan pada proses klasifikasi. Beberapa kata yang memiliki bobot TF-IDF tertinggi ditunjukkan pada Tabel 2

Tabel 2. Kata dengan Bobot TF-IDF Tertinggi

Kata	Bobot
ai	147,43
yg	66,55
manusia	49,26

Kata	Bobot
belajar	40,30
jadi	39,64
bang	38,64
orang	36,16
apa	34,08
sama	32,15
aja	31,07

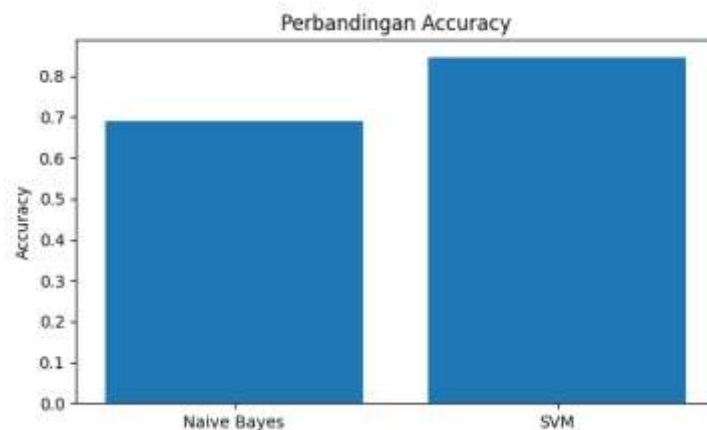
Hasil ekstraksi fitur menunjukkan bahwa kata "ai" menjadi kata yang paling dominan dalam diskusi pengguna media sosial. Selain itu, kata-kata seperti "manusia", "belajar", "orang", dan "teknologi" juga sering muncul. Temuan ini mengindikasikan bahwa diskusi masyarakat lebih banyak berfokus pada hubungan antara perkembangan teknologi AI dengan kemampuan manusia dalam menghadapi perubahan dunia kerja.

3.2. Pengujian dan Perbandingan Kinerja Algoritma

Pengujian dilakukan menggunakan dua algoritma klasifikasi yaitu Naïve Bayes dan Support Vector Machine (SVM). Dataset dibagi menjadi data training sebanyak 1.892 data dan data testing sebanyak 473 data dengan perbandingan 80:20. Evaluasi model dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score.

Tabel 3. Hasil Pengujian Algoritma

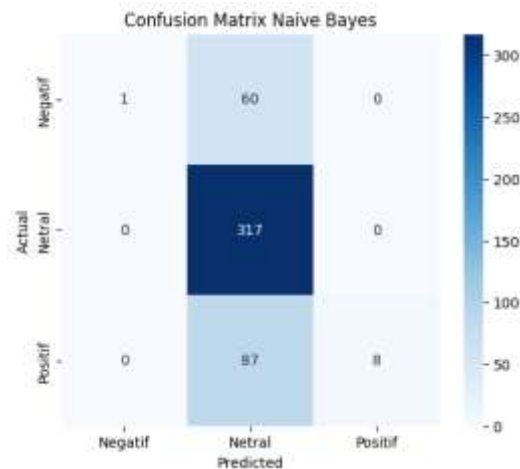
Metode	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	68,92%	78,77%	68,92%	57,94%
SVM	84,57%	87,09%	84,57%	82,93%



Gambar 2. Perbandingan Accuracy Naïve Bayes dan SVM

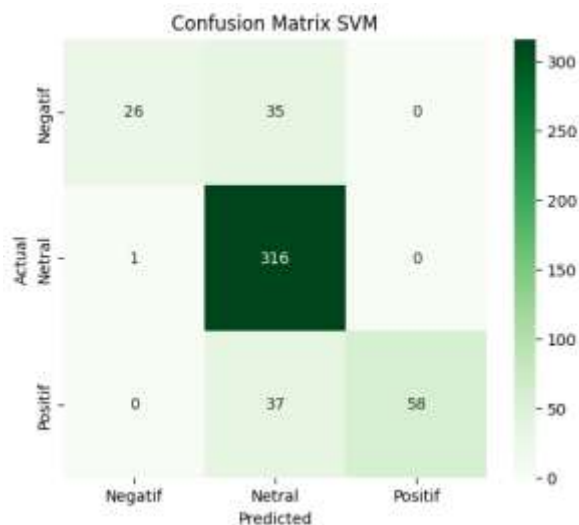
Berdasarkan hasil pengujian diperoleh bahwa algoritma SVM menghasilkan performa yang lebih baik dibandingkan Naïve Bayes pada seluruh metrik evaluasi. SVM memperoleh accuracy sebesar 84,57%, sedangkan Naïve Bayes hanya memperoleh accuracy sebesar 68,92%. Selisih accuracy antara kedua algoritma mencapai 15,65%.

Perbedaan performa juga terlihat pada nilai F1-score. Algoritma SVM memperoleh F1-score sebesar 82,93%, sedangkan Naïve Bayes hanya mencapai 57,94%. Hasil tersebut menunjukkan bahwa SVM memiliki kemampuan yang lebih baik dalam mengenali seluruh kelas sentimen secara seimbang.



Gambar 3. Confusion matrix Naïve Bayes

Berdasarkan confusion matrix Naïve Bayes terlihat bahwa model cenderung mengklasifikasikan sebagian besar data ke dalam kelas netral. Dari 61 data sentimen negatif, hanya satu data yang berhasil diklasifikasikan dengan benar. Kondisi serupa juga terjadi pada kelas positif, dimana hanya delapan data yang berhasil dikenali secara tepat. Hasil tersebut menunjukkan bahwa Naïve Bayes mengalami kesulitan dalam mengidentifikasi kelas minoritas akibat distribusi data yang tidak seimbang.



Gambar 4. Confusion matrix SVM

Berbeda dengan Naïve Bayes, algoritma SVM menunjukkan kemampuan klasifikasi yang lebih baik. Model berhasil mengidentifikasi 26 data sentimen negatif dan 58 data sentimen positif secara benar. Pada kelas netral, model mampu mengklasifikasikan 316 dari 317 data dengan tepat. Hasil ini menunjukkan bahwa SVM mampu membangun batas pemisah yang lebih optimal pada data teks hasil ekstraksi TF-IDF.

3.3 Pembahasan

Hasil penelitian menunjukkan bahwa sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence dalam dunia kerja didominasi oleh sentimen netral. Temuan ini mengindikasikan bahwa sebagian besar masyarakat

belum memiliki pandangan yang sepenuhnya positif maupun negatif terhadap perkembangan AI. Banyak pengguna yang memanfaatkan kolom komentar untuk berdiskusi, bertanya, dan memberikan opini terkait dampak AI terhadap pekerjaan di masa depan.

Dominasi kata-kata seperti "ai", "manusia", dan "belajar" menunjukkan bahwa perhatian utama masyarakat terletak pada kemampuan manusia untuk beradaptasi terhadap perkembangan teknologi. Sebagian pengguna memandang AI sebagai peluang untuk meningkatkan produktivitas dan efisiensi kerja, sementara sebagian lainnya mengungkapkan kekhawatiran mengenai kemungkinan berkurangnya lapangan pekerjaan akibat otomatisasi.

Dari sisi metode klasifikasi, hasil penelitian membuktikan bahwa Support Vector Machine lebih unggul dibandingkan Naïve Bayes. Keunggulan tersebut terlihat dari nilai accuracy, precision, recall, dan F1-score yang lebih tinggi. Kemampuan SVM dalam menangani data berdimensi tinggi menjadikannya lebih efektif dalam melakukan klasifikasi teks yang telah direpresentasikan menggunakan TF-IDF. Sebaliknya, asumsi independensi antar fitur yang digunakan oleh Naïve Bayes menyebabkan performanya menurun ketika dihadapkan pada data teks yang kompleks dan distribusi kelas yang tidak seimbang.

Berdasarkan keseluruhan hasil penelitian dapat disimpulkan bahwa Support Vector Machine merupakan algoritma yang lebih efektif untuk melakukan analisis sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence dalam dunia kerja. Struktur ini biasanya jauh lebih disukai dosen pembimbing dan reviewer karena lebih ringkas, mengalir, dan tidak terkesan seperti laporan praktikum.

4. Kesimpulan

Penelitian ini berhasil melakukan analisis sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence (AI) dalam dunia kerja menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM). Data penelitian diperoleh melalui proses scraping komentar YouTube yang menghasilkan sebanyak 2.365 komentar. Setelah melalui tahapan preprocessing, pelabelan sentimen, dan ekstraksi fitur menggunakan TF-IDF, data kemudian diklasifikasikan menggunakan kedua algoritma tersebut. Hasil analisis sentimen menunjukkan bahwa mayoritas komentar pengguna berada pada kategori netral sebanyak 1.617 komentar atau 68,37%. Sentimen positif berjumlah 474 komentar atau 20,04%, sedangkan sentimen negatif berjumlah 274 komentar atau 11,59%. Temuan ini menunjukkan bahwa sebagian besar pengguna media sosial masih memberikan tanggapan yang bersifat informatif, diskusi, maupun pertanyaan terkait perkembangan AI dalam dunia kerja tanpa menunjukkan kecenderungan yang kuat terhadap dukungan maupun penolakan. Berdasarkan hasil pengujian model, algoritma Naïve Bayes memperoleh nilai accuracy sebesar 68,92%, precision sebesar 78,77%, recall sebesar 68,92%, dan F1-score sebesar 57,94%. Sementara itu, algoritma Support Vector Machine memperoleh nilai accuracy sebesar 84,57%, precision sebesar 87,09%, recall sebesar 84,57%, dan F1-score sebesar 82,93%. Hasil penelitian menunjukkan bahwa Support Vector Machine memiliki performa yang lebih baik dibandingkan Naïve Bayes dalam melakukan klasifikasi sentimen pengguna media sosial terhadap penggunaan Artificial Intelligence dalam dunia kerja. Keunggulan tersebut ditunjukkan oleh nilai accuracy dan F1-score yang lebih tinggi pada seluruh metrik evaluasi. Dengan demikian, Support Vector Machine dapat direkomendasikan sebagai metode yang lebih efektif untuk digunakan dalam analisis sentimen pada data teks yang memiliki karakteristik serupa dengan penelitian ini.

Reference

- [1] Y. K. Dwivedi *et al.*, "International Journal of Information Management Opinion Paper: ' So what if ChatGPT wrote it? ' Multidisciplinary perspectives on opportunities , challenges and implications of generative conversational AI for research , practice and policy ☆," vol. 71, no. March, 2023, doi: 10.1016/j.ijinfomgt.2023.102642.
- [2] G. Damioli, V. Van Roy, and D. Vertesy, "The impact of artificial intelligence on labor productivity," *Eurasian Bus. Rev.*, vol. 11, no. 1, pp. 1–25, 2021, doi: 10.1007/s40821-020-00172-8.
- [3] D. Rizki, S. Pratama, T. A. Munandar, and K. Fadhilla, "Multinomial Naive Bayes Algorithm for Indonesian language Sentiment Classification Related to Jakarta International Stadium (JIS)," vol. 2, no. 1, pp. 12–22, 2024.
- [4] A. Gerliandeva, Y. H. Chrisnanto, and H. Ashaury, "Optimasi Klasifikasi Sentimen pada Komentar Online menggunakan Multinomial Naive Bayes dan Ekstraksi Fitur TF-IDF serta N-grams Optimization of Sentiment Classification on Online Comments using Multinomial Naive Bayes and TF-IDF Feature Extraction and N-grams," no. X, pp. 259–272, 2024, doi: 10.56873/jpkm.v9i2.5585.
- [5] M. Iqbal and E. Budianto, "Model Predictive Analysis of Performance in Training and Course Institutions Using Naive Bayes and K-Means Clustering," pp. 2703–2708, 2025.
- [6] H. Mustakim and S. Priyanta, "Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 16, no. 2, p. 113, 2022, doi: 10.22146/ijccs.68903.
- [7] K. T. Aprilia, I. P. Sitorus, M. R. Ridha, and M. S. Novelan, "Comparison of Random Forest and Naïve Bayes Classifier Methods for Monkeypox Classification," vol. 3, no. 1, pp. 58–67, 2026.

DOI: <https://doi.org/10.31004/riggs.v5i2.12229>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

- [8] S. M. Sadid, J. C. Young, and A. Rusli, "Spam Filtering on User Feedback via Text Classification using Multinomial Naïve Bayes and TF-IDF," vol. 13, no. 2, pp. 108–113, 2021.
- [9] S. Darma, A. Jihad, A. Fayed, S. M. P. Pardede, and M. H. Aqsha, "Predictive Analysis of Flood Risk Factors Based on a Machine Learning Approach : Comparative Study of SVM and XGBoost Algorithms," vol. 3, no. 1, pp. 24–33, 2026.
- [10] H. Afrizal, Z. Sitorus, and M. S. Novelan, "Fault Detection And Recovery System On 20 Kv Distribution Network Using Real-Time Analysis With Support Vector Machine Algorithm," vol. 18, 2026.
- [11] Z. S. Muhammad Ridho Pradana, Rio Septian Hardinata, "ANALISA ALGORITMA SUPPORT VECTOR MACHINE PADA DATA BUNGA IRIS," vol. 30, no. April, pp. 477–487, 2022.
- [12] M. S. Novelan and M. Iqbal, "Comparative Analysis of the Performance of K-Nearest Neighbor (K-NN) and Naive Bayes Algorithms on User Satisfaction Levels of the Tokopedia Application conduct a comparative performance evaluation of," 2025.
- [13] A. Fatma and P. Indah, "Analisis Sentimen Komentar YouTube MV K-Pop Menggunakan Naïve Bayes dan SVM : Studi Kasus Jung Jaehyun ‘ Horizon ;,'" vol. 15, no. 2, pp. 383–395, 2025.
- [14] A. Kurniawan and A. P. Widyassari, "Komparasi Naïve Bayes dan Support Vector Machine dalam Klasifikasi Sentimen Pelantikan Menteri Purbaya Melalui YouTube," vol. 5, pp. 762–767, 2026.
- [15] T. D. Permana, "PERBANDINGAN PERFORMA ALGORITMA NAIVE BAYES DAN TERHADAP INDUSTRI ESPORTS DI INDONESIA," vol. 2, no. 6, pp. 1391–1399, 2025.