

Terbit online pada laman : <http://teknosi.fti.unand.ac.id/>

Jurnal Nasional Teknologi dan Sistem Informasi

| ISSN (Print) 2460-3465 | ISSN (Online) 2476-8812 |



Artikel Penelitian

Analisis Sentimen Opini Publik Terkait Kebijakan Royalti Musik Menggunakan Algoritma Naïve Bayes pada Data Twitter

Nilna Muna Sholihah ^{a,*}, Mas Rahmah ^b

^{ab}Program Studi Magister Pengembangan Sumber Daya Manusia, Sekolah Pascasarjana, Universitas Airlangga, Surabaya 60286

Informasi Artikel

Sejarah Artikel:

Diterima Redaksi: 14 November 2025

Revisi Akhir: 18 Desember 2025

Diterbitkan Online: 23 Agustus 2026

Kata Kunci

Analisis Sentimen,
Naïve Bayes,
Orange Data Mining,
Royalti Musik,
Twitter

Korespondensi

E-mail: nilnmn001@gmail.com *

ABSTRACT

Kebijakan royalti musik di Indonesia yang diatur melalui Peraturan Menteri Hukum Nomor 27 Tahun 2025 telah memicu beragam respons dari masyarakat. Penelitian ini bertujuan untuk menganalisis sentimen opini publik di media sosial Twitter terkait kebijakan royalti musik menggunakan algoritma Naïve Bayes. Data dikumpulkan dari Twitter berbahasa Indonesia pada periode 1-15 Agustus 2025, menghasilkan 1.588 tweet yang kemudian dianalisis menggunakan Orange Data Mining. Proses analisis meliputi *crawling data* menggunakan Google Colab, *preprocessing data*, analisis emosional, dan analisis polarisasi menggunakan *multilingual sentiment*. Hasil penelitian menunjukkan distribusi sentimen, yaitu 584 tweet positif (36,8%), 561 tweet netral (35,3%), dan 443 tweet negatif (27,9%). Algoritma Naïve Bayes menghasilkan akurasi sebesar 97,17% dengan tingkat kesalahan prediksi sebanyak 45 tweet. *Confusion matrix* menunjukkan performa model yang sangat baik dalam mengklasifikasikan sentimen publik. Penelitian ini memberikan wawasan penting mengenai persepsi masyarakat terhadap implementasi kebijakan royalti musik dan dapat menjadi bahan evaluasi bagi pemangku kepentingan dalam industri musik Indonesia.

1. PENDAHULUAN

Industri musik Indonesia mengalami perkembangan signifikan dalam beberapa tahun terakhir, sejalan dengan digitalisasi dan pertumbuhan *platform streaming* musik. Namun, isu mengenai hak kekayaan intelektual, khususnya royalti musik, masih menjadi perdebatan yang kompleks di tengah masyarakat. Peraturan Menteri Hukum Republik Indonesia Nomor 27 Tahun 2025 tentang Pengelolaan Royalti Hak Cipta Lagu dan/atau Musik merupakan implementasi dari Peraturan Pemerintah Nomor 56 Tahun 2021 dan Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta, yang mengatur mekanisme pembayaran royalti musik di Indonesia [1].

Data dari Direktorat Jenderal Kekayaan Intelektual (DJKI) menunjukkan pertumbuhan signifikan dalam penerimaan royalti musik di Indonesia. Pada tahun 2020, royalti yang diterima dari 14 sektor pengguna jasa musik mencapai Rp29,1 miliar,

kemudian mengalami penurunan menjadi Rp19,9 miliar pada 2021 akibat dampak pandemi COVID-19. Namun, tren positif kembali muncul dengan penerimaan royalti mencapai Rp35 miliar pada 2022 dan melonjak menjadi Rp55,2 miliar pada 2023, dengan target penerimaan sebesar Rp120 miliar pada tahun 2024 [2]. Meskipun terdapat peningkatan dalam penerimaan royalti, potensi penghasilan pencipta lagu Indonesia diperkirakan dapat mencapai Rp4 triliun pada tahun 2023, namun realisasinya masih jauh dari harapan [3].

Kebijakan royalti musik ini mewajibkan setiap pengguna komersial yang memutar musik di ruang publik untuk membayar royalti melalui Lembaga Manajemen Kolektif Nasional (LMKN). Implementasi kebijakan ini bertujuan untuk melindungi hak pencipta, musisi, dan pemegang hak cipta lainnya agar mendapatkan kompensasi yang adil atas karya mereka. Di beberapa negara seperti Amerika Serikat, Jepang, Inggris, dan

Korea Selatan, sistem serupa sudah diberlakukan sejak lama sebagai bentuk perlindungan hak kekayaan intelektual [4]. Namun, kebijakan ini juga menuai pro dan kontra di kalangan pelaku usaha, terutama sektor usaha mikro, kecil, dan menengah (UMKM), yang merasa terbebani dengan kewajiban pembayaran royalti.

Isu royalti musik yang mulai ramai diperbincangkan pada pertengahan tahun 2024 menciptakan polarisasi opini di kalangan masyarakat. Di satu sisi, para musisi dan pencipta lagu melihat kebijakan ini sebagai bentuk pengakuan dan perlindungan hak ekonomi mereka di era *streaming* digital. Sabina dan Ramli [5] dalam penelitiannya mengungkapkan bahwa era *streaming* musik digital telah mengubah pola konsumsi dan distribusi karya musik secara global, termasuk di Indonesia, yang membawa peluang sekaligus tantangan dalam perlindungan hak cipta. Pasar musik digital Indonesia diproyeksikan mencapai US\$363,30 juta pada tahun 2024 menurut data Statista, menunjukkan potensi ekonomi yang signifikan bagi industri musik nasional [6]. Namun, tantangan transparansi royalti, penegakan hukum lintas yurisdiksi, dan pembajakan digital masih menjadi permasalahan yang menghambat implementasi kebijakan yang efektif. Di sisi lain, pelaku UMKM menghadapi tantangan dalam memenuhi kewajiban pembayaran royalti akibat kondisi ekonomi yang belum stabil pasca-pandemi, kurangnya sosialisasi, dan ketidakjelasan mekanisme pembayaran yang kerap menimbulkan resistensi dari pelaku usaha kecil.

Kompleksitas industri musik digital juga melibatkan berbagai aktor baru yang belum sepenuhnya diatur dalam regulasi hak cipta yang ada. Wiratama dan Baidhowi [7] menjelaskan bahwa ketiadaan pengaturan eksplisit terhadap agregator dalam Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta menyebabkan kekosongan tanggung jawab hukum dalam hubungan antara pencipta, agregator, DJKI, dan *digital service provider* (DSP). Riset Koalisi Seni pada tahun 2023 menemukan bahwa kebijakan hak cipta di Indonesia masih berbasis fisik-analog dan belum meladeni perkembangan dalam industri musik digital, sehingga posisi musisi menjadi semakin rentan [8]. Tullah dkk. [9] menegaskan bahwa meskipun UU Hak Cipta telah memberikan perlindungan hukum berupa hak moral dan hak ekonomi terhadap pencipta, namun implementasinya dalam konteks digital masih menghadapi berbagai kendala, terutama terkait transparansi distribusi royalti dan mekanisme pengelolaan oleh Lembaga Manajemen Kolektif.

Perkembangan media sosial, khususnya Twitter (kini X), telah menjadi ruang publik digital yang signifikan bagi masyarakat Indonesia dalam mengekspresikan opini dan berdiskusi tentang berbagai isu. Data menunjukkan bahwa Indonesia memiliki 25,25 juta pengguna Twitter per Juli 2023, menempatkan Indonesia di peringkat ke-4 dunia dalam jumlah pengguna Twitter [10]. Pada Januari 2025, jumlah pengguna X (Twitter) di Indonesia mencapai 25,2 juta pengguna dengan jangkauan iklan mencapai 8,8% dari total populasi [11]. Platform ini menjadi wadah bagi masyarakat untuk mengekspresikan pendapat, kritik, dan dukungan terhadap berbagai isu publik, termasuk kebijakan royalti musik.

Peran media sosial dalam membentuk opini publik terhadap kebijakan pemerintah telah menjadi fenomena yang semakin

signifikan di Indonesia. Aryantini dkk. [12] menjelaskan bahwa media sosial telah menjadi ruang publik digital yang memungkinkan masyarakat untuk mengekspresikan pendapat, mengakses informasi, dan berpartisipasi dalam diskusi kebijakan secara masif dan *real-time*. Dalam konteks kebijakan royalti musik, media sosial menjadi arena penting dimana berbagai perspektif dari musisi, pelaku UMKM, akademisi, dan masyarakat umum saling bertemu dan berdebat. Saputra dkk. [13] mengidentifikasi fenomena "*viral based policy*" di Indonesia, dimana pemerintah seringkali baru merespons isu setelah menjadi viral di media sosial tanpa perencanaan yang matang, yang dapat menimbulkan hambatan pada proses implementasi kebijakan. Analisis sentimen terhadap opini publik di Twitter menjadi penting karena dapat memberikan gambaran komprehensif tentang persepsi masyarakat secara *real-time*, yang dapat menjadi masukan bagi pembuat kebijakan dalam mengevaluasi dan menyempurnakan implementasi kebijakan royalti musik di Indonesia [14].

Twitter sebagai sumber data opini publik telah banyak dimanfaatkan dalam berbagai penelitian analisis sentimen. Penelitian Fitri dkk. [15] menggunakan Twitter untuk menganalisis sentimen kampanye anti-LGBT di Indonesia dengan akurasi 86,43% menggunakan algoritma Naïve Bayes. Studi lain oleh Riyantoro dan Fauziah [16] menganalisis sentimen Twitter mengenai Pilkada Indonesia dan berhasil mengklasifikasikan sentimen publik dengan tingkat akurasi yang baik. Penelitian tentang pandemi COVID-19 juga memanfaatkan Twitter sebagai sumber data, seperti yang dilakukan oleh Damanik dan Setyohadi [17] yang menganalisis sentimen publik tentang COVID-19 di Indonesia menggunakan metode Multinomial Naïve Bayes dan Support Vector Machine.

Analisis sentimen merupakan metode yang efektif untuk mengekstraksi dan mengklasifikasikan opini publik menjadi kategori positif, negatif, atau netral. Algoritma Naïve Bayes telah terbukti menjadi salah satu metode yang paling efektif untuk klasifikasi sentimen pada data Twitter berbahasa Indonesia. Penelitian oleh Syafutra dan Kusri [18] menunjukkan bahwa Naïve Bayes berhasil mengklasifikasikan sentimen publik terhadap Omnibus Law dengan hasil yang memuaskan. Studi terbaru oleh Setiawati dkk. [19] menganalisis 576.764 tweet tentang pengangguran di Indonesia selama dan pasca COVID-19 menggunakan Naïve Bayes dan Support Vector Machine, menunjukkan efektivitas metode ini dalam menganalisis data Twitter skala besar.

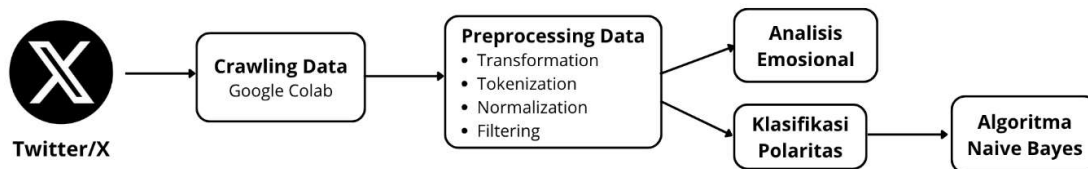
Konteks kebijakan royalti musik di Indonesia yang masih relatif baru dan kontroversial, ditambah dengan besarnya populasi pengguna Twitter di Indonesia, menjadikan analisis sentimen Twitter sebagai pendekatan yang relevan untuk memahami persepsi publik. Penelitian ini penting untuk memberikan gambaran komprehensif tentang opini masyarakat yang dapat menjadi masukan bagi pembuat kebijakan dalam mengevaluasi dan menyempurnakan implementasi kebijakan royalti musik di Indonesia. Dengan demikian, penelitian ini diharapkan dapat menjembatani kesenjangan antara kebijakan pemerintah, kepentingan pelaku industri kreatif, dan kebutuhan pelaku usaha, khususnya UMKM.

2. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen terhadap data Twitter. Pendekatan kuantitatif dipilih karena memungkinkan analisis data dalam jumlah besar secara objektif dan terukur melalui teknik komputasional [20]. Penelitian dilakukan secara sistematis melalui beberapa tahapan seperti pada gambar 1, yaitu pengumpulan data, *preprocessing*, analisis emosional, analisis polarisasi, dan klasifikasi menggunakan algoritma Naïve Bayes. Kerangka kerja sistematis ini mengikuti metodologi standar dalam penelitian *text mining* dan *sentiment analysis* sebagaimana direkomendasikan oleh Liu [21] dalam analisis opini berbasis teks.

2.1. Crawling Data

Pada tahap awal, data penelitian dikumpulkan dari platform Twitter menggunakan teknik *crawling* dengan Google Colab.



Gambar 1. Tahapan Penelitian

2.2. Pre-processing Data

Setelah mendapatkan data, dilanjutkan ke tahap *preprocessing data* menggunakan Orange Data Mining. Orange Data Mining merupakan *platform open-source* berbasis *visual programming* yang menyediakan komponen lengkap untuk *text mining* dan *machine learning* [24]. Tahap *preprocessing* bertujuan untuk mentransformasi data teks mentah menjadi format yang sesuai untuk analisis komputasional dengan menghilangkan *noise* dan inkonsistensi [25]. Tahap ini digunakan untuk menghapus URL, *mention* (@username), *hashtag*, *emoticon*, dan karakter khusus yang tidak relevan.

Dalam tahapan ini meliputi *transformation*, *tokenization*, *normalization*, dan *filtering*. *Transformation* dilakukan untuk mengubah semua teks menjadi huruf kecil (*lowercase*) guna memastikan konsistensi dalam pemrosesan teks [26]. *Tokenization* adalah proses memecah teks menjadi unit-unit terkecil atau token, biasanya berupa kata individual, yang merupakan langkah fundamental dalam *natural language processing* [27]. *Normalization* melibatkan proses *stemming* untuk mengubah kata-kata ke bentuk dasarnya menggunakan algoritma Sastrawi yang telah dioptimalkan untuk bahasa Indonesia [28]. *Filtering* menggunakan *stopword* untuk menghapus kata-kata umum yang tidak memiliki makna signifikan, seperti “yang”, “di”, “ke”, “dari”, “dan”, dll. Proses *preprocessing* ini sangat penting untuk meningkatkan kualitas data dan mengurangi *noise* yang dapat mempengaruhi akurasi model klasifikasi.

2.3. Analisis Emosional

Setelah *preprocessing* selesai, analisis emosional dilakukan untuk mengidentifikasi emosi yang terkandung dalam tweet, seperti senang (*joy*), sedih (*sadness*), marah (*anger*), takut (*fear*),

Web crawling merupakan teknik otomatis untuk mengekstraksi data dari *website* atau *platform* media sosial secara sistematis dan terstruktur [22]. Penggunaan Google Colab dipilih karena menyediakan lingkungan Python berbasis *cloud* dengan akses gratis ke sumber daya komputasi yang memadai untuk pemrosesan data skala besar [23]. Periode pengumpulan data dilakukan pada tanggal 1-15 Agustus 2025. Pemilihan periode ini didasarkan pada beberapa pertimbangan, dimana periode tersebut merupakan masa awal implementasi Peraturan Menteri Hukum Nomor 27 Tahun 2025. Selain itu, tingginya diskusi publik di media sosial mengenai kebijakan royalti musik pada periode tersebut, serta terdapat sosialisasi intensif dari pemerintah terkait kewajiban pembayaran royalti musik. Kriteria data yang dikumpulkan meliputi tweet berbahasa Indonesia, menggunakan kata kunci terkait “royalti musik”, dan bukan retweet otomatis untuk menghindari duplikasi.

jijik (*disgust*), dan terkejut (*surprise*). Hal ini didasarkan pada konsep 6 dasar emosional manusia oleh Paul Ekman [29]. Teori Ekman menyatakan bahwa enam emosi dasar ini bersifat universal dan dapat diidentifikasi melalui ekspresi wajah maupun bahasa tekstual di berbagai budaya [30]. Dalam Orange Data Mining, algoritma yang digunakan adalah Ekman (*multiclass*) untuk menganalisis emosi dari tweet yang telah didapatkan.

2.4. Klasifikasi Polaritas

Selanjutnya data *preprocessing* dilanjutkan dengan analisis polarisasi menggunakan *multilingual sentiment* untuk mengklasifikasikan sentimen tweet ke dalam tiga kategori, yaitu positif, negatif, dan netral. Klasifikasi polaritas merupakan tahap fundamental dalam *sentiment analysis* yang bertujuan mengidentifikasi orientasi opini secara keseluruhan [31]. Metode ini memanfaatkan model *pre-trained* yang telah dilatih pada dataset multibahasa termasuk Bahasa Indonesia. Penggunaan model *pre-trained multilingual* memiliki keunggulan dalam mengatasi keterbatasan data *training* berbahasa Indonesia dan memanfaatkan *transfer learning* dari bahasa-bahasa lain yang memiliki karakteristik linguistik serupa [32]. Model *multilingual* telah terbukti efektif dalam analisis sentimen untuk *low-resource languages* seperti bahasa Indonesia dengan performa yang kompetitif [33].

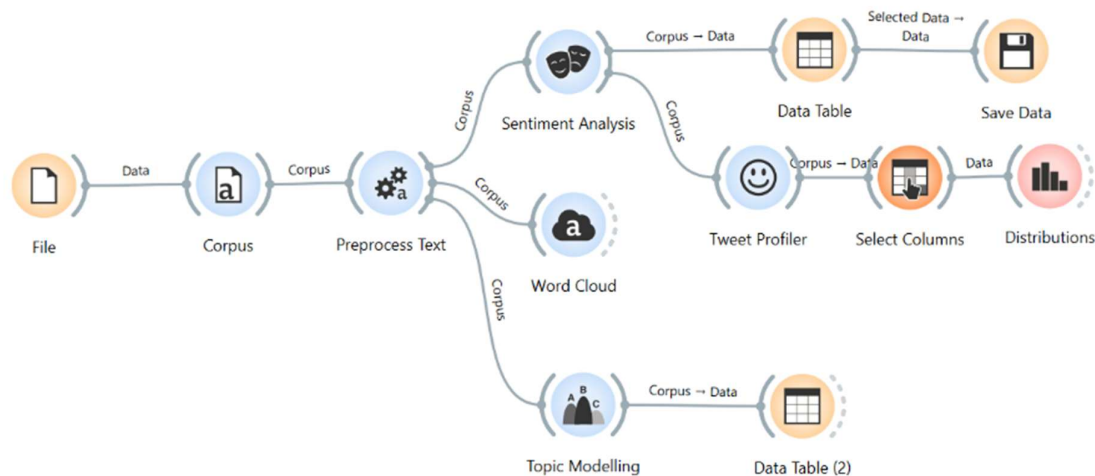
2.5. Analisis Naive Bayes

Data yang dihasilkan dari klasifikasi polarisasi diberikan label dan diproses menggunakan algoritma Naïve Bayes. Naïve Bayes merupakan algoritma *machine learning* berbasis probabilitas yang mengimplementasikan teorema Bayes dengan asumsi independensi kondisional antar fitur [34]. Meskipun asumsi independensi ini jarang terpenuhi dalam praktik, Naïve Bayes tetap menunjukkan performa yang sangat baik dalam berbagai aplikasi *text classification*, termasuk *sentiment analysis* [35].

Dataset dibagi menjadi 80% data *training* dan 20% data *testing*. Model *training* digunakan untuk melatih model Naïve Bayes, sedangkan model *testing* untuk menguji model pada data *testing*.

Selanjutnya, evaluasi model menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*. *Confusion matrix* menyajikan visualisasi detail tentang performa klasifikasi dengan menampilkan jumlah *true positive*, *true negative*, *false positive*,

dan *false negative* untuk setiap kelas [36]. Metrik *accuracy* mengukur proporsi prediksi yang benar secara keseluruhan, *precision* mengukur ketepatan prediksi positif, *recall* mengukur kemampuan model menemukan semua *instance* positif, sedangkan *F1-score* merupakan *harmonic mean* dari *precision* dan *recall* yang memberikan gambaran performa seimbang [37].



Gambar 2. Model Analisis Sentimen

3. HASIL

3.1. Crawling Data

Proses *crawling data* Twitter menggunakan Google Colab pada periode 1-15 Agustus 2025 menghasilkan 1.588 tweet berbahasa Indonesia yang membahas tentang kebijakan royalti musik. Dataset ini mencakup tweet original dan *reply* yang mengandung kata kunci relevan dengan topik penelitian, yaitu “royalti musik”. Data yang terkumpul menunjukkan bahwa isu royalti musik menjadi topik yang cukup hangat diperbincangkan di media sosial Twitter selama periode tersebut.

3.2. Pre-processing Data

Selanjutnya, dilakukan *preprocessing data* menggunakan Online Data Mining untuk menghapus URL, *mention* (@username), *hashtag*, *emoticon*, dan karakter khusus yang tidak relevan. Tahapan ini meliputi *transformation*, *tokenization*, *normalization*, dan *filtering*. *Filtering* menggunakan *stopword* untuk menghapus kata-kata umum yang tidak memiliki makna signifikan, seperti “yang”, “di”, “ke”, “dari”, “dan”, dll. Pada gambar 3, dapat diketahui bagaimana gambaran *word cloud* sebelum dilakukan *preprocessing data*. Dataset mentah mengandung banyak *noise* seperti URL, *mention*, *hashtag*, *emoticon*, dan karakter khusus. Setelah melalui tahap *preprocessing*, dihasilkan gambaran *word cloud* seperti pada gambar 4. Data menjadi lebih bersih dan terstruktur dengan hanya mempertahankan kata-kata bermakna. Proses *preprocessing* ini dapat mengurangi dimensi data dan menghilangkan *noise* yang tidak relevan, sehingga meningkatkan efektivitas analisis sentimen.



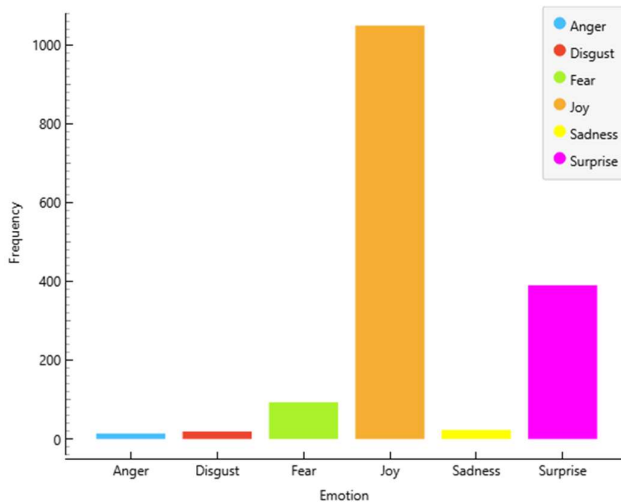
Gambar 3. Sebelum *Preprocessing Data*



Gambar 4. Sesudah *Preprocessing Data*

3.3. Analisis Emosional

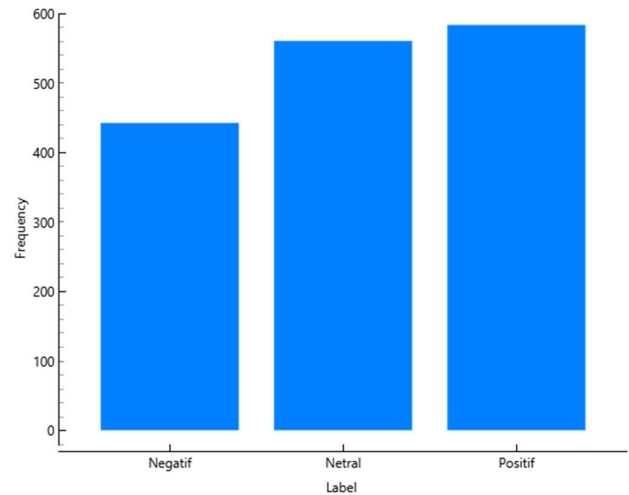
Setelah data lebih bersih, dilakukan analisis emosional untuk mengidentifikasi emosi yang terkandung dalam tweet, seperti senang (*joy*), sedih (*sadness*), marah (*anger*), takut (*fear*), jijik (*disgust*), dan terkejut (*surprise*). Algoritma yang digunakan adalah Ekman (*multiclass*) untuk menganalisis emosi dari tweet yang telah didapatkan. Hasil analisis pada gambar 5 menunjukkan ekspresi marah (*anger*) = 14 (0,88%), jijik (*disgust*) = 19 (1,20%), takut (*fear*) = 93 (5,86%), senang (*joy*) = 1.049 (66,06%), sedih (*sadness*) = 23 (1,45%), dan terkejut (*surprise*) = 390 (24,56%).



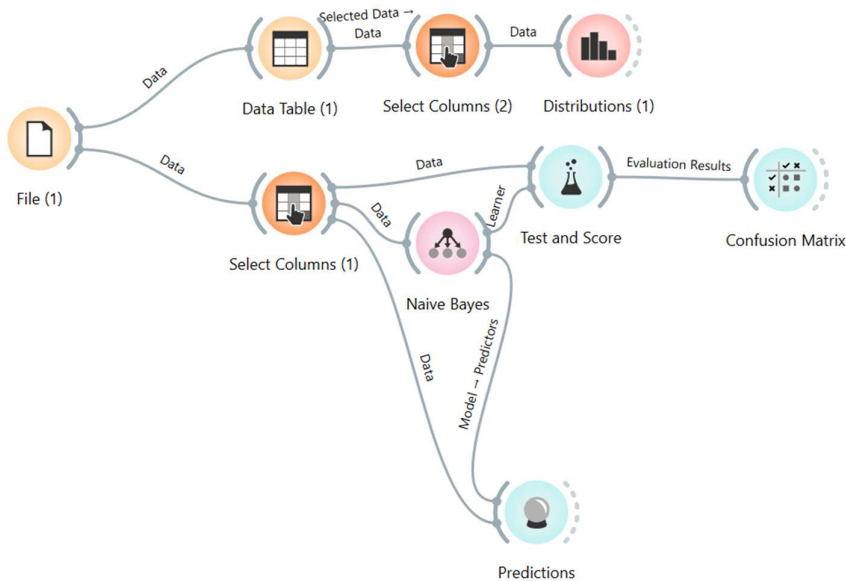
Gambar 5. Hasil Analisis Emosional

3.4. Klasifikasi Polarisasi

Selain analisis emosional, dilakukan klasifikasi polarisasi menggunakan *multilingual sentiment* untuk mengklasifikasikan sentimen tweet ke dalam tiga kategori, yaitu positif, negatif, dan netral. Pada gambar 6, hasil analisis menunjukkan bahwa dari 1588 data, sentimen netral memiliki angka maksimum sebesar 561 tweet (35,33%), sentimen positif sebesar 584 tweet (36,78%), dan sentimen negatif sebesar 443 tweet (27,90%).



Gambar 6. Hasil Klasifikasi Polarisasi



Gambar 7. Model Analisis Naïve Bayes

3.5. Analisis Naïve Bayes

Hasil data klasifikasi sentimen selanjutnya diproses menggunakan algoritma Naïve Bayes dengan model analisis pada gambar 7. Dataset dibagi menjadi 80% data *training* dan 20% data *testing*. Model *training* digunakan untuk melatih model Naïve Bayes, sedangkan model *testing* untuk menguji model pada data *testing*. Selanjutnya, evaluasi model menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*.

3.5.1. Test and Score

Pada gambar 8, hasil *test and score* menunjukkan evaluasi model menggunakan *cross-validation 10-fold* yang menghasilkan metrik evaluasi, seperti *accuracy* sebesar 97,17% yang menunjukkan proporsi prediksi yang benar dari total prediksi, *precision* sebesar 97,23% yang menunjukkan proporsi prediksi positif yang benar, *recall* sebesar 97,17% yang menunjukkan proporsi data positif aktual yang teridentifikasi dengan benar, *F1-Score* sebesar 97,19% yang menunjukkan harmonik *mean* dari

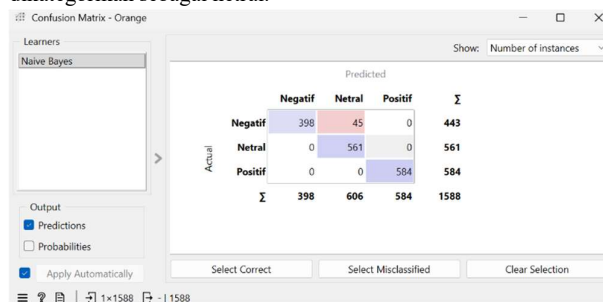
precision dan *recall*, serta AUC sebesar 0,986 yang menunjukkan kemampuan model membedakan kelas dengan sangat baik. Hasil ini menunjukkan bahwa algoritma Naïve Bayes memiliki performa yang sangat baik dalam mengklasifikasikan sentimen tweet mengenai kebijakan royalti musik. Akurasi di atas 97% mengindikasikan bahwa model dapat diandalkan untuk mengidentifikasi sentimen opini publik.



Gambar 8. Test and Score Naïve Bayes

3.5.2. Confusion Matrix

Selanjutnya, *confusion matrix* memberikan visualisasi detail tentang performa model dalam mengklasifikasikan setiap kategori sentimen. Dari 1588 data, instansi yang diklasifikasikan dengan benar sejumlah 1543, sedangkan 45 instansi salah diklasifikasikan (*misclassified*). Pada kelas negatif, total 443 instansi aktual negatif, dimana 398 berhasil diprediksi dengan benar sebagai negatif dan 45 salah diprediksi sebagai netral. Pada kelas netral, total 561 instansi aktual netral yang juga berhasil diprediksi dengan benar sebagai netral. Pada kelas positif, total 584 instansi aktual positif yang juga berhasil diprediksi dengan benar sebagai positif. Model Naïve Bayes menunjukkan performa yang sangat baik, walaupun hanya sedikit kesulitan dalam mengklasifikasikan beberapa instansi negatif, yang keliru dikategorikan sebagai netral.



Gambar 9. Confusion Matrix Naïve Bayes

4. PEMBAHASAN

Hasil penelitian pada analisis emosional menunjukkan bahwa mayoritas tweet didominasi oleh emosi senang (*joy*) sebesar 66,06% (1.049 tweet), diikuti oleh terkejut (*surprise*) sebesar 24,56% (390 tweet). Emosi positif yang sangat dominan ini mengindikasikan bahwa sebagian besar diskusi publik di Twitter mengenai kebijakan royalti musik memiliki sentimen yang cenderung mendukung atau setidaknya menerima kebijakan tersebut dengan baik. Emosi senang yang tinggi ini sejalan dengan data distribusi royalti LMKN yang menunjukkan peningkatan signifikan dari Rp27,8 miliar pada 2022 menjadi Rp54,2 miliar pada 2024 [38], yang mencerminkan optimisme kalangan musisi dan pencipta lagu terhadap perlindungan hak

ekonomi mereka. Sementara itu, emosi negatif seperti takut (*fear*) mencapai 5,86% (93 tweet), sedih (*sadness*) 1,45% (23 tweet), dan marah (*anger*) 0,88% (14 tweet), dengan total emosi negatif sebesar 8,19%. Meskipun proporsi emosi negatif relatif kecil, keberadaan emosi takut yang mencapai hampir 6% sejalan dengan paparan Kepala Program Studi Manajemen Fakultas Ekonomi dan Bisnis UGM, I Wayan Nuka Lantara, PhD [39], yang mengungkapkan kekhawatiran pelaku usaha mikro dan kecil terhadap beban pembayaran royalti, terutama dalam kondisi ekonomi yang belum stabil pasca-pandemi. Emosi jijik (*disgust*) sebesar 1,20% (19 tweet) kemungkinan merefleksikan ketidakpuasan terhadap mekanisme implementasi kebijakan atau persepsi ketidakadilan dalam penerapan tarif royalti.

Hasil analisis sentimen terhadap 1.588 tweet mengenai kebijakan royalti musik menunjukkan distribusi yang relatif seimbang dengan sentimen positif mencapai 36,8% (584 tweet), netral 35,3% (561 tweet), dan negatif 27,9% (443 tweet). Distribusi sentimen yang hampir merata ini mengindikasikan adanya polarisasi opini publik terhadap kebijakan tersebut. Fenomena polarisasi di media sosial bukanlah hal baru dalam konteks kebijakan publik di Indonesia. Penelitian Rahmawati [40] menjelaskan bahwa algoritma media sosial berpotensi menciptakan ruang gema (*echo chamber*) yang memperkuat bias konfirmasi dan memperparah polarisasi opini di kalangan pengguna. Dalam konteks kebijakan royalti musik, polarisasi ini mencerminkan dua perspektif yang berbeda, yaitu perlindungan hak ekonomi kreator musik versus kekhawatiran beban ekonomi pelaku UMKM. Studi oleh Annas dkk. [41] menegaskan bahwa media sosial telah menjadi arena penting dalam pembentukan opini publik terhadap kebijakan pemerintah, dimana diskusi dapat berlangsung secara cepat namun juga rentan terhadap penyebaran informasi yang tidak lengkap atau menyesatkan.

Tweet dengan sentimen positif umumnya berasal dari kalangan musisi, pencipta lagu, dan pendukung hak kekayaan intelektual. Mereka mengapresiasi kebijakan ini sebagai bentuk perlindungan hak cipta dan kesejahteraan kreator musik. Tweet netral sebagian besar berisi informasi faktual, pertanyaan mengenai mekanisme pembayaran, atau diskusi objektif tentang kebijakan. Tweet ini tidak menunjukkan sikap pro atau kontra yang jelas. Sedangkan untuk tweet dengan sentimen negatif banyak berasal dari pelaku usaha kecil, pemilik kafe, dan UMKM yang merasa terbebani. Distribusi sentimen yang relatif seimbang ini menunjukkan bahwa kebijakan royalti musik masih menjadi isu kontroversial dengan pendapat yang terpolarisasi.

Performa algoritma Naïve Bayes dalam penelitian ini mencapai akurasi 97,17% dengan *precision* 97,23%, *recall* 97,17%, dan *F1-score* 97,19%, yang menunjukkan hasil yang sangat memuaskan dalam mengklasifikasikan sentimen tweet berbahasa Indonesia. Hasil ini secara signifikan lebih tinggi dibandingkan dengan penelitian sejenis. Damayanti dan Lhaksana [42] dalam analisis sentimen Pilpres 2024 di Twitter menggunakan Naïve Bayes hanya mencapai akurasi 64%, sementara Sembiring dan Dewa [43] mencapai akurasi 72,19% dalam analisis sentimen pemilihan umum. Perbedaan akurasi yang cukup besar ini dapat dijelaskan oleh beberapa faktor metodologis. Pertama, penelitian ini menggunakan *preprocessing* yang lebih komprehensif dengan tahapan *cleansing*, *case folding*, *tokenization*, *stopword removal*, dan *stemming* menggunakan algoritma Sastrawi yang telah

dioptimalkan untuk Bahasa Indonesia. Kedua, penggunaan *multilingual analysis* untuk pelabelan awal menghasilkan data *training* yang lebih akurat dibandingkan pelabelan manual yang rentan terhadap subjektivitas. Ketiga, fokus topik penelitian yang spesifik pada satu isu kebijakan menghasilkan dataset yang lebih homogen dengan karakteristik linguistik yang lebih konsisten, sehingga memudahkan model dalam melakukan klasifikasi.

Confusion matrix penelitian ini menunjukkan performa klasifikasi yang sangat baik dengan total kesalahan prediksi (*misclassified*) hanya 45 tweet dari 1.588 tweet (2,83%). Pada kelas negatif, total 443 instansi aktual negatif, dimana 398 berhasil diprediksi dengan benar sebagai negatif dan 45 salah diprediksi sebagai netral. Pada kelas netral, total 561 instansi aktual netral yang juga berhasil diprediksi dengan benar sebagai netral. Pada kelas positif, total 584 instansi aktual positif yang juga berhasil diprediksi dengan benar sebagai positif. Model Naïve Bayes menunjukkan performa yang sangat baik, walaupun hanya sedikit kesulitan dalam mengklasifikasikan beberapa instansi negatif, yang keliru dikategorikan sebagai netral. Adapun kesalahan terbanyak terjadi pada klasifikasi antara sentimen netral dan positif, yang dapat dipahami mengingat ambiguitas *inherent* dalam bahasa natural, terutama pada tweet yang bersifat informatif namun memiliki nuansa positif.

Hasil ini konsisten dengan temuan Lin dan Nuha [44] yang mengembangkan strategi *hybrid deep-learning* untuk analisis sentimen dataset Indonesia dan menemukan bahwa klasifikasi antara kategori netral dan positif merupakan tantangan tersendiri karena batas yang tipis antara kedua kategori tersebut. Perbandingan dengan penelitian lain menunjukkan bahwa meskipun beberapa studi menggunakan metode yang lebih canggih seperti Support Vector Machine atau *deep learning*, Naïve Bayes tetap menunjukkan performa yang kompetitif bahkan superior dalam konteks tertentu. Rosyida dkk. [45] menemukan bahwa SVM mencapai akurasi 76,33% dalam analisis sentimen Pilpres 2024, masih jauh di bawah akurasi Naïve Bayes dalam penelitian ini.

Implikasi praktis dari penelitian ini menunjukkan perlunya pendekatan komunikasi yang lebih strategis dari pemerintah dalam mensosialisasikan kebijakan royalti musik. Polarisasi opini yang teridentifikasi melalui analisis sentimen mengindikasikan bahwa kebijakan masih memerlukan penyesuaian, terutama terkait mekanisme keringanan untuk UMKM sebagaimana dijelaskan oleh DJKI. Fenomena polarisasi di media sosial, sebagaimana dianalisis oleh Purboyo dkk. [46], dapat menyebabkan eskalasi partisanship yang memperburuk perpecahan sosial jika tidak dimitigasi dengan baik. Oleh karena itu, pemerintah perlu meningkatkan transparansi dalam pengelolaan dan distribusi royalti melalui LMKN, menyederhanakan prosedur pembayaran, dan menyediakan skema yang lebih fleksibel bagi pelaku usaha mikro dan kecil. Metodologi analisis sentimen menggunakan Naïve Bayes yang terbukti efektif dalam penelitian ini juga dapat diadopsi oleh pembuat kebijakan untuk monitoring dan evaluasi berkelanjutan terhadap respon publik di media sosial, sehingga kebijakan dapat disesuaikan secara responsif berdasarkan *feedback real-time* dari masyarakat.

Hasil penelitian ini memiliki beberapa keterbatasan, antara lain data hanya dikumpulkan dari Twitter, sehingga tidak mencerminkan opini publik secara keseluruhan. Selain itu, periode pengumpulan data yang relatif pendek (15 hari), serta analisis terbatas pada teks berbahasa Indonesia formal, sehingga tweet dengan bahasa gaul atau campur bahasa mungkin tidak teranalisis optimal.

5. KESIMPULAN

Berdasarkan hasil penelitian analisis sentimen opini publik mengenai kebijakan royalti musik menggunakan algoritma Naïve Bayes pada data Twitter, dapat disimpulkan beberapa pada distribusi sentimen: Opini publik terhadap kebijakan royalti musik menunjukkan distribusi yang relatif seimbang dengan sentimen positif (36,8%), netral (35,3%), dan negatif (27,9%). Hal ini mengindikasikan bahwa kebijakan masih menjadi isu kontroversial yang memerlukan sosialisasi dan penyesuaian lebih lanjut.

Performa Model: Algoritma Naïve Bayes menunjukkan performa yang sangat baik dalam mengklasifikasikan sentimen tweet berbahasa Indonesia dengan akurasi mencapai 97,17%, *precision* 97,23%, *recall* 97,17%, F1-score 97,19%, dan AUC 0,986. Hasil ini membuktikan bahwa Naïve Bayes merupakan algoritma yang efektif untuk analisis sentimen pada data Twitter.

Analisis Emosional: Emosi dominan yang muncul dalam diskusi publik adalah *joy* (senang), mencakup 66,06%. Emosi dominan kedua adalah *surprise* (terkejut) dengan 24,56%. Kedua emosi ini secara kolektif merepresentasikan mayoritas sentimen non-negatif yang sangat kuat (sekitar 90,62%). Meskipun persentase emosi lainnya kecil, keberadaan emosi negatif ini tetap menunjukkan adanya kekhawatiran atau keberatan spesifik dalam sebagian kecil segmen masyarakat yang perlu diperhatikan.

Kesalahan Prediksi: Dari 1.588 tweet, terdapat 45 kesalahan prediksi, dengan kesalahan terbanyak terjadi pada klasifikasi antara sentimen netral dan negatif. Hal ini wajar mengingat ambiguitas batas antara kedua kategori tersebut dalam konteks bahasa natural.

DAFTAR PUSTAKA

- [1] Kementerian Hukum, “Peraturan Pelaksanaan Peraturan Pemerintah Nomor 56 Tahun 2021 tentang Pengelolaan Royalti Hak Cipta Lagu dan/atau Musik,” peraturan.bpk.go.id.
- [2] A. Setiawan, “Royalti Perkuat Industri Musik Nasional,” indonesia.go.id.
- [3] Henry, “Potensi Penghasilan Pencipta Lagu Indonesia dari Royalti Selama 2023 Capai Rp4 Triliun, Faktanya Jauh dari Harapan,” liputan6.com.
- [4] IWM, “Mempelajari Kewajiban Bayar Royalti untuk Bisnis Non-Musik,” Direktorat Jenderal Kekayaan Intelektual.
- [5] R. Sabina dan T. Ramli, “Analisis Pelindungan Hak Cipta di Era Streaming Musik Digital: Implementasi dan tantangan di Indonesia,” *Causa: Jurnal Hukum dan Kewarganegaraan*, vol. 11, no. 5, 2025.
- [6] N. Yusuf, “Complete Guide to Music Copyright Laws in Indonesia,” Am Badar & Am Badar.
- [7] R. Wiratama dan B. Baidhowi, “Tanggung Jawab Hukum Aggregator dalam Distribusi Musik tanpa Pendaftaran Hak

- Cipta di Indonesia,” *Nomos: Jurnal Penelitian Ilmu Hukum*, vol. 5, no. 4, 2025.
- [8] A. Ikhlasanti, “Siaran Pers: Riset Hak Cipta Musik Digital: Kebijakan Belum Memihak Seniman,” *Koalisi Seni*.
 - [9] D. Tullah, H. Khalid, dan S. Poernomo, “Perlindungan Hak Cipta Dalam Industri Musik Digital di Indonesia Dalam Perspektif Hukum Perdata,” *Journal of Lex Philosophy (JLP)*, vol. 5, no. 2, 2024.
 - [10] C. M. Annur, “Jumlah Pengguna Twitter Indonesia Duduki Peringkat ke-4 Dunia per Juli 2023,” *databoks.co.id*.
 - [11] S. Kemp, “Digital 2025: Indonesia,” *datareportal.com*.
 - [12] P. Aryantini, G. Pramana, T. Erviantono, dan E. Duarte, “Peran Media Sosial Dalam Pembentukan Opini Publik Politik: Studi Kasus Kampanye Pemilu 2024,” *Triwikrama: Jurnal Ilmu Sosial*, vol. 5, no. 5, 2024.
 - [13] A. Saputra, D. Kurniasih, dan D. Indranika, “Dinamika Kebijakan Publik di Indonesia dalam Menghadapi Isu Viral: Antara Respons dan Tantangan Implementasi,” *Jurnal Administrasi Publik*, vol. 16, no. 1, 2025.
 - [14] K. Latifah dan F. Najicha, “Implikasi Media Sosial Terhadap Formulasi Kebijakan Publik,” *Jurnal Kewarganegaraan*, vol. 6, no. 1, 2023.
 - [15] V. A. Fitri, R. Andreswari, dan M. A. Hasibuan, “Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm,” *Procedia Comput Sci*, vol. 161, pp. 765–772, 2019, doi: [10.1016/j.procs.2019.11.181](https://doi.org/10.1016/j.procs.2019.11.181).
 - [16] Riyantoro dan Fauziah, “Sentiment Analysis of Twitter Data on Indonesia’s Cabinet Using Naïve Bayes and Support Vector Machine Algorithms,” *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 30, no. 2, pp. 108–118, Aug. 2025, doi: [10.35760/tr.2025.v30i2.13953](https://doi.org/10.35760/tr.2025.v30i2.13953).
 - [17] F. J. Damanik dan D. B. Setyohadi, “Analysis of Public Sentiment About Covid-19 in Indonesia on Twitter Using Multinomial Naive Bayes and Support Vector Machine,” *IOP Conf Ser Earth Environ Sci*, vol. 704, no. 1, p. 012027, Mar. 2021, doi: [10.1088/1755-1315/704/1/012027](https://doi.org/10.1088/1755-1315/704/1/012027).
 - [18] A. D. Syafutra dan K. Kusriani, “Analisis Sentimen Omnibus Law di Twitter dengan Machine Learning dan Teknik Resampling,” *Jambura Journal of Informatics*, vol. 1, no. 1, pp. 22–35, Apr. 2025, doi: [10.37905/jji.v1i1.30935](https://doi.org/10.37905/jji.v1i1.30935).
 - [19] P. A. Setiawati, I. M. A. D. Suarjaya, dan I. N. P. Trisna, “Sentiment Analysis of Unemployment in Indonesia During and Post COVID-19 on X (Twitter) Using Naïve Bayes and Support Vector Machine,” *Journal of Information Systems and Informatics*, vol. 6, no. 2, pp. 662–675, Jun. 2024, doi: [10.51519/journalisi.v6i2.713](https://doi.org/10.51519/journalisi.v6i2.713).
 - [20] J. W. Creswell dan J. D. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, Fifth edition. United States of America: SAGE Publications, 2018.
 - [21] B. Liu, “Sentiment Analysis and Opinion Mining,” *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, May 2012, doi: [10.2200/S00416ED1V01Y201204HLT016](https://doi.org/10.2200/S00416ED1V01Y201204HLT016).
 - [22] B. Zhao, “Web Scraping,” in *Encyclopedia of Big Data*, Cham: Springer International Publishing, 2017, pp. 1–3. doi: [10.1007/978-3-319-32001-4_483-1](https://doi.org/10.1007/978-3-319-32001-4_483-1).
 - [23] T. Carneiro, R. V. Medeiros Da Nobrega, T. Nepomuceno, G.-B. Bian, V. H. C. De Albuquerque, dan P. P. R. Filho, “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications,” *IEEE Access*, vol. 6, pp. 61677–61685, 2018, doi: [10.1109/ACCESS.2018.2874767](https://doi.org/10.1109/ACCESS.2018.2874767).
 - [24] J. Demsar *et al.*, “Orange: Data mining toolbox in Python,” *Journal of Machine Learning Research*, vol. 14, no. 1, 2013.
 - [25] S. Vijayanari, J. Ilamathi, dan Nithya, “Preprocessing Techniques for Text Mining-An Overview,” *International Journal of Computer Science & Communication Networks*, vol. 5, no. 1, 2015.
 - [26] A. Jivani, “A Comparative Study of Stemming Algorithms,” *International Journal of Computer Technology and Applications*, vol. 2, no. 6, 2011.
 - [27] J. J. Webster dan C. Kit, “Tokenization as the Initial Phase in NLP,” in *Proceedings of the 14th conference on Computational linguistics*, Morristown, NJ, USA: Association for Computational Linguistics, 1992, p. 1106. doi: [10.3115/992424.992434](https://doi.org/10.3115/992424.992434).
 - [28] M. Adriani, J. Asian, B. Nazief, S. M. M. Tahaghoghi, dan H. E. Williams, “Stemming Indonesian,” *ACM Transactions on Asian Language Information Processing*, vol. 6, no. 4, pp. 1–33, Dec. 2007, doi: [10.1145/1316457.1316459](https://doi.org/10.1145/1316457.1316459).
 - [29] T. Y. Pahtoni dan H. Jati, “Analisis Sentimen Data Twitter Terkait CHATGPT Menggunakan Orange Data Mining,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 2, pp. 329–336, Apr. 2024, doi: [10.25126/jtiik.20241127276](https://doi.org/10.25126/jtiik.20241127276).
 - [30] P. Ekman dan D. Cordaro, “What is Meant by Calling Emotions Basic,” *Emotion Review*, vol. 3, no. 4, pp. 364–370, Oct. 2011, doi: [10.1177/1754073911410740](https://doi.org/10.1177/1754073911410740).
 - [31] B. Pang dan L. Lee, “Opinion Mining and Sentiment Analysis,” *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008, doi: [10.1561/15000000011](https://doi.org/10.1561/15000000011).
 - [32] J. Devlin, M.-W. Chang, K. Lee, dan K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proceedings of NAACL-HLT*, 2019.
 - [33] S. Cahyawijaya *et al.*, “IndoNLG: Benchmark and Resources for Evaluating Indonesian Natural Language Generation,” in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2021.
 - [34] C. D. Manning, P. Raghavan, dan H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008. doi: [10.1017/CBO9780511809071](https://doi.org/10.1017/CBO9780511809071).
 - [35] I. Rish, “An Empirical Study of the Naive Bayes Classifier,” *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, vol. 3, no. 22, 2001.
 - [36] D. Powers, “Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation,” *Journal of Machine Learning Technologies*, vol. 2, no. 1, 2011.
 - [37] M. Sokolova dan G. Lapalme, “A Systematic Analysis of Performance Measures for Classification Tasks,” *Inf Process Manag*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: [10.1016/j.ipm.2009.03.002](https://doi.org/10.1016/j.ipm.2009.03.002).
 - [38] “Pelantikan LMKN: Royalti Aman, Kreator Nyaman,” *Kanwil Kemenkum Yogyakarta*.

- [39] N. Rosa, "Pakar UGM: Kebijakan Royalti Musik Kurang Tepat, Ekonomi Sedang Tidak Stabil," *detik.com*.
- [40] D. Rahmawati, "Risiko Polarisasi Algoritma Media Sosial: Kajian Terhadap Kerentanan Sosial dan Ketahanan Bangsa, Media Sosial dan Dampaknya bagi Sosial Budaya Masyarakat Indonesia," *Jurnal Kajian Lemhanas RI*, vol. 33, 2018.
- [41] F. B. Annas, H. N. Petranto, dan A. A. Pramayoga, "Opini publik dalam polarisasi politik di media sosial," *Jurnal PIKOM (Penelitian Komunikasi dan Pembangunan)*, vol. 20, no. 2, p. 111, Dec. 2019, doi: [10.31346/jpikom.v20i2.2006](https://doi.org/10.31346/jpikom.v20i2.2006).
- [42] L. Damayanti dan K. M. Lhaksmana, "Sentiment Analysis of the 2024 Indonesia Presidential Election on Twitter," *Sinkron*, vol. 8, no. 2, pp. 938–946, Mar. 2024, doi: [10.33395/sinkron.v8i2.13379](https://doi.org/10.33395/sinkron.v8i2.13379).
- [43] A. R. Sembiring dan C. K. Dewa, "Sentiment Analysis On Indonesian Tweets about the 2024 Election," *Sinkron*, vol. 9, no. 1, pp. 413–422, Jan. 2025, doi: [10.33395/sinkron.v9i1.14481](https://doi.org/10.33395/sinkron.v9i1.14481).
- [44] C.-H. Lin dan U. Nuha, "Sentiment Analysis of Indonesian Datasets Based on a Hybrid Deep-Learning Strategy," *J Big Data*, vol. 10, no. 1, p. 88, May 2023, doi: [10.1186/s40537-023-00782-9](https://doi.org/10.1186/s40537-023-00782-9).
- [45] Tamara Rosyida, H. P. Putro, dan Herry Wahyono, "Analisis Sentimen Terhadap Pilpres 2024 Berdasarkan Opini dari Twitter Menggunakan Naïve Bayes dan SVM," *TEKNOKRIS*, vol. 26, no. 1, pp. 23–32, Jun. 2023, doi: [10.61488/teknokris.v26i1.245](https://doi.org/10.61488/teknokris.v26i1.245).
- [46] M. G. Purboyo, J. Junaidi, dan A. Amriwan, "Eskalasi Partisanship Politik dalam Pembentukan Polarisasi pada Masyarakat Digital," *SOSIOLOGI: Jurnal Ilmiah Kajian Ilmu Sosial dan Budaya*, vol. 26, no. 01, pp. 27–45, Mar. 2024, doi: [10.23960/sosiologi.v26i01.1220](https://doi.org/10.23960/sosiologi.v26i01.1220).