Analisis Sentimen Publik Terhadap Produk Rekomendasi Influencer Menggunakan Naive Bayes: Studi Kasus Tasya Farasya

Garda Khoerunnisa[1)*], Syti Sarah Maesaroh[2)], Muhammad Rizki Nugraha[3]
[1,2,3)] Universitas Pendidikan Indonesia

Received: 23 June 2025 Accepted: 27 June 2025 Published: 1 July 2025

*
gardakhoerunnisa06@gmail.com

Kata Kunci: Analisis Sentimen, Semi-Supervised Learning, TikTok, Naive Bayes, Tasya Farasya Approved

DSI: Jurnal Data Science Indonesia is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak : Media sosial telah mengubah pola konsumsi informasi dan menjadikan *influencer* sebagai rujukan utama, khususnya dalam industri kecantikan. Penelitian ini bertujuan untuk menelaah persepsi publik terhadap produk Wardah yang berlabel Tasya Farasya *Approved*, serta menguji efektivitas metode *Semi-Supervised Learning* (SSL) dalam analisis sentimen komentar di TikTok. Sebanyak 2.716 komentar berhasil dikumpulkan melalui *web scraping* dari video milik Tasya Farasya dan akun resmi Wardah selama tahun 2024–2025. Setelah melalui tahap pelabelan dan *preprocessing*, diperoleh 878 komentar yang kemudian dibagi menjadi 80% data berlabel yang digunakan sebagai data latih dan data uji, serta 20% data tidak berlabel yang digunakan dalam proses *pseudo-labeling*. Seluruh data direpresentasikan menggunakan metode TF-IDF dan diklasifikasikan menggunakan algoritma *Naive Bayes*. Model divalidasi menggunakan *5-fold-cross-validation* dan diseimbangkan dengan teknik SMOTE. SSL diterapkan melalui metode *pseudo-labeling* terhadap data tidak berlabel dengan tingkat kepercayaan ≥ 80%. Hasil akhir menunjukkan distribusi sentimen terdiri dari 425 komentar positif, 305 netral, dan 148 negatif. Sentimen positif didominasi oleh apresiasi terhadap produk micellar water. Pengaruh Tasya Farasya sebagai influencer turut memperkuat persepsi positif publik. Penerapan SSL meningkatkan akurasi model dari 0,7565 menjadi 0,7733, disertai peningkatan *precision*, *recall*, dan *F1-score*.

PENDAHULUAN

Seiring pesatnya perkembangan teknologi digital, media sosial menjadi ruang interaksi utama masyarakat. Rata-rata waktu penggunaan media sosial mencapai 3 jam 11 menit per hari [1], dengan TikTok menempati posisi teratas sebagai aplikasi yang paling sering digunakan [2]. Meningkatnya durasi konsumsi konten digital mendorong popularitas *influencer* yang memainkan peran penting dalam membentuk pandangan dan opini publik [2]. Di industri kecantikan, salah satu *influencer* berpengaruh adalah Tasya Farasya, yang dikenal lewat konten ulasan dan rekomendasi produk. Label "Tasya Farasya *Approved*" pada produk menjadi daya tarik tersendiri karena dianggap mencerminkan kualitas unggul [3].

Namun, belum banyak penelitian yang secara langsung menelaah persepsi publik terhadap produk berlabel tersebut melalui komentar di media sosial. Metode yang sebelumnya digunakan umumnya berupa survei atau studi literatur [4][5], ang memiliki keterbatasan dalam menangkap opini publik secara alami [6]. Karena itu, diperlukan pendekatan alternatif seperti analisis sentimen berbasis big data media sosial. Analisis sentimen memungkinkan pengolahan data teks tidak terstruktur untuk menggali opini publik [7]. Pendekatannya meliputi *lexicon-based*, *machine learning*, dan *hybrid* [8], di mana metode TF-IDF dan *Naive Bayes* terbukti efektif [4].

Untuk mengatasi keterbatasan data berlabel, digunakan *semi-supervised learning* (SSL) dengan

pseudo-labeling [9]. enelitian ini menjadi yang pertama mengaplikasikan metode SSL untuk analisis sentimen terhadap komentar TikTok mengenai produk berlabel "Tasya Farasya Approved". Tujuan dari penelitian ini adalah untuk menganalisis persepsi publik, memetakan distribusi sentimen, serta mengevaluasi efektivitas pendekatan SSL dalam meningkatkan performa klasifikasi sentimen.

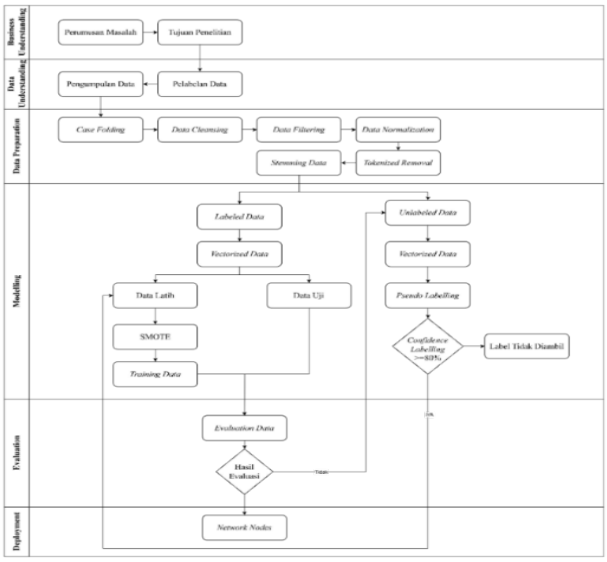
TINJAUAN LITERATUR

Analisis sentimen bertujuan untuk mengolah data teks tidak terstruktur guna memahami opini subjektif [7]. Metode ini banyak digunakan dalam pemasaran, evaluasi produk, dan media sosial karena kemampuannya menangkap opini publik secara terstruktur [7]. Salah satu pendekatan umum dalam analisis sentimen adalah machine learning, seperti algoritma *Naive Bayes* yang memanfaatkan *Teorema Bayes* dengan asumsi independensi antar fitur [10]. Seiring perkembangannya, *semi-supervised learning* (SSL) menjadi alternatif yang kian populer karena dapat menggabungkan data berlabel dan tidak berlabel secara efisien [11]. Beberapa teknik SSL yang sering digunakan antara lain *self-training*, *label propagation*, dan *co-training*. Dalam praktiknya, TF-IDF digunakan untuk memberi bobot pada kata berdasarkan frekuensinya dalam korpus [12].

Penelitian ini juga menyoroti peran *influencer* dalam komunikasi digital. Tasya Farasya, misalnya, memiliki pengaruh besar di industri kecantikan lewat konten edukatif dan jujur, dengan engagement rate TikTok mencapai 8,92% [13][14]. Label "Tasya Farasya *Approved*" bahkan dianggap sebagai indikator kualitas yang memengaruhi keputusan pembelian [15]. TikTok sendiri dinilai efektif dalam promosi digital karena didukung algoritma rekomendasi dan interaktivitas pengguna yang tinggi [16][17]. Penelitian ini menggunakan pendekatan *Cross Industry Standard Process for Data Mining* (CRISP-DM) yang mencakup enam tahap: pemahaman bisnis, data, persiapan data, pemodelan, evaluasi, dan penerapan [18], termasuk *text preprocessing seperti tokenisasi*, *stemming*, *stopword removal*, dan pelabelan data [19][20].

Hingga kini, belum ada penelitian yang secara langsung menelaah persepsi publik terhadap produk berlabel Tasya Farasya *Approved* menggunakan komentar dari media sosial. Sebagian besar studi sebelumnya mengandalkan survei atau studi literatur [4][5], yang memiliki keterbatasan dalam menangkap opini publik secara spontan dan mendalam [6]. Oleh karena itu, penelitian ini bertujuan untuk menganalisis persepsi publik terhadap produk Wardah melalui komentar TikTok dengan pendekatan SSL. Selain itu, penelitian ini juga menguji efektivitas SSL dalam klasifikasi sentimen pada data terbatas, serta memperkaya literatur mengenai pengaruh label influencer dalam membentuk persepsi konsumen di media sosial berbasis video pendek.

METODE PENELITIAN



Gambar 1. Desain Penelitian

Dari gambar 1 dijelaskan mengenai alur dari proses penelitian yang teridiri dari beberapa tahap yaitu:

1. *Business Understanding* : Tahap ini berfokus pada pemahaman tujuan penelitian yang ingin dicapai guna menghasilkan insight yang relevan. Proses ini mencakup identifikasi permasalahan, penentuan tujuan analisis, serta perancangan strategi yang dapat menjawab kebutuhan penelitian secara efektif.
2. *Data Understanding* : Tahap ini berfokus pada pengumpulan data menggunakan teknik web scraping. Proses pengambilan data dilakukan secara otomatis melalui skrip Python yang telah dimodifikasi dari kode sumber yang tersedia di GitHub (Topneyz, 2024).
3. *Data Preparation* : Tahap ini berfokus untuk mempersiapkan data sebelum memasuki tahap analisis lebih lanjut. Proses ini meliputi beberapa tahapan yaitu *transform case*, *Cleaning*, *Filtering*, *Normalization Data*, *Tokenize*, *Stopwords*, dan *Stemming*.
4. *Modelling* : Pendekatan modelling dalam penelitian ini mengadaptasi metode Farhan dkk. (2023) dengan penyesuaian. Data dibagi menjadi data berlabel (80%) dan tidak berlabel (20%). Model awal dilatih menggunakan 5-fold cross-validation dan SMOTE pada data berlabel yang telah dikonversi ke bentuk numerik menggunakan TF-IDF. Algoritma Naive Bayes digunakan untuk pelatihan, dengan evaluasi menggunakan akurasi, classification report, dan confusion matrix. Selanjutnya, pendekatan semi-supervised learning diterapkan melalui metode pseudo-labeling, di mana hanya prediksi dengan tingkat kepercayaan ≥80% yang dijadikan pseudo-labeled dan digabungkan ke data latih awal untuk pelatihan ulang guna meningkatkan performa model.
5. *Evaluation* : Pada tahap ini, dilakukan proses evaluasi untuk mengukur performa model dalam mengklasifikasikan data. Evaluasi dilakukan dengan menggunakan data uji yang berbeda dari data yang dipakai saat pelatihan, dengan membagi *labeled data* menjadi 5 bagian dimana 4 bagian digunakan untuk melatih model dan bagian lainnya digunakan menjadi data uji. Beberapa metrik yang digunakan dalam evaluasi meliputi akurasi, *classification report* dan *confusion matrix*.
6. *Deployment* : Hasil utama penelitian disajikan dalam bentuk visualisasi *Network Nodes* yang menggambarkan hubungan antar kata atau istilah paling sering muncul dalam komentar pengguna untuk masing-masing kategori sentimen, yaitu positif, negatif, dan netral. Visualisasi ini membantu mengungkap pola dan keterkaitan kata-kata kunci yang menjadi ciri khas dari setiap sentimen, sehingga memberikan gambaran yang lebih mendalam tentang bagaimana masyarakat memandang produk Wardah yang direkomendasikan oleh Tasya Farasya.

HASIL PENELITIAN

1. *Data Scrapping*

Data yang berhasil dikumpulkan melalui proses scraping mencakup tiga kolom, yaitu nomor indeks, Video ID, dan komentar. Pengambilan data dilakukan dari akun TikTok milik Tasya Farasya dan Wardah, dengan rentang waktu publikasi video mulai dari tahun 2024 hingga 2025. Rentang waktu ini dipilih agar data yang diperoleh bersifat terkini dan relevan dengan tren saat ini. Proses pengumpulan data dilakukan menggunakan bahasa pemrograman Python 3.10, dijalankan di lingkungan Visual Studio Code. Skrip scraping memanfaatkan pustaka requests untuk mengakses endpoint komentar TikTok secara langsung dalam format JSON, serta pustaka csv untuk menyimpan data dalam format tabular. Skrip dirancang untuk mengekstrak komentar dari beberapa video TikTok secara otomatis, dengan identifikasi Video ID dari URL, iterasi komentar berdasarkan cursor, dan penyimpanan hasil scraping ke dalam file dataset.csv. Proses ini menghasilkan total 2.716 komentar yang digunakan sebagai data utama dalam analisis sentimen yang dapat dilihat pada Gambar 2.

index	No	Video ID	Comment
0	1	7453072689318726917	Aku pake yg glow cushion wardah, bagus bgt😍
1	2	7453072689318726917	plisss aku keringetan parahhh ini bakal longsor ga🙏
2	3	7453072689318726917	Mending mana ya wardah colorfit atau instaperfect?
3	4	7453072689318726917	Buat aku si kulit kering, cocok bgt pake perfect glow cushion ini plisss😍
4	5	7453072689318726917	aku ada yg ijooo, oksidasi banget truss berattt 😭 agak nyesel belinya, krna udah exited denger claimnya cocok buat oily.. 😩
5	6	7453072689318726917	Wardah Colorfit Cushion emang pliss cakep banget 🍒🍒🍒aku pake yang glow pake yang cloud bagus banget 😭
6	7	7453072689318726917	Kalau buat oilly skin bagusan yang mana kaa
7	8	7453072689318726917	ada yg mau ga cushion dari wardah colorfit yg ijo krn shade ny ga cocok di aku shade 22n
8	9	7453072689318726917	Wardah color fit longsor g kk kalau buat kulit normal to oily
9	10	7453072689318726917	guys, aku preloved cushion wardah colorfit perfect glow shade 23W. pemakaiannya baru 2x aku buka harga di 90rb bisa nego ya, ada yang mau?
10	11	7453072689318726917	bingunh 22n apa 21w, yg mirip kyk skintint you rose vanilla yg mana pls
11	12	7453072689318726917	pake coushen color fit bikin abu-abu
12	13	7453072689318726917	ada yg mau gantiin cushion Wardah colorfit glow shade 32n natural beige gak?masih baru jual murah
13	14	7453072689318726917	udah nyobain juga yg hijau cocok untuk kulit ku super oily Banget
14	15	7453072689318726917	kulit oily mending wardah/make over?
15	16	7453072689318726917	pake yg glow cocok sih tp shade 32n skg kegelapan k aku tp 22n keterangan 😭 kumaha nyak ?
16	17	7453072689318726917	wardah ijo sm skintific biru bgus mana?
17	18	7453072689318726917	ka oily bagus wardah colorfit hijau atau bnb pink
18	19	7453072689318726917	ada yg mau ngga cushion wardah colorfit yg glow blum pernah dipake salah shade,shade 32N 60k aja

Gambar 2. Hasil *Scrapping Data*

2. Analisis Sentimen Menggunakan Metode CRISP-DM

2.1 *Business Understanding*

Pada tahap *business understanding* dalam penelitian ini, dilakukan analisis sentimen dengan tujuan untuk mengelompokkan sentimen pengguna TikTok ke dalam tiga kategori, yaitu positif, negatif, dan netral. Analisis ini menggunakan algoritma *Naive Bayes* untuk mengetahui bagaimana sentimen masyarakat terhadap produk Wardah, yang merupakan salah satu produk rekomendasi dari Tasya Farasya.

2.2 *Data Understanding*

Pada tahap ini, dilakukan proses pemilihan atribut berupa komentar yang akan digunakan dalam pelatihan. Hanya komentar yang memiliki makna atau isi yang dipilih, sementara komentar yang terdiri dari satu kata atau hanya berisi emotikon tidak disertakan. Selanjutnya, data yang telah terpilih diberi label secara manual menggunakan perangkat lunak *Microsoft Excel*. Proses labelling data dalam penelitian ini bertujuan untuk mengkategorikan komentar pengguna ke dalam tiga kategori sentimen, yaitu positif yang disimbolkan dengan 2, netral dengan simbol 1, dan negatif dengan 0 sebagai simbolnya. Data yang telah diberi label ini digunakan sebagai dataset awal untuk keperluan pelatihan model klasifikasi sentimen. Proses pelabelan dilakukan sebelum data dibersihkan agar konteks asli dari komentar tetap terjaga dan makna sentimen tidak hilang akibat proses normalisasi teks. Pendekatan ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa pelabelan sebelum praproses dapat mempertahankan keaslian ekspresi sentimen, sehingga meningkatkan akurasi dalam proses klasifikasi. Pelabelan dilakukan oleh peneliti sendiri dengan mengacu pada pedoman sederhana berdasarkan isi komentar, seperti penggunaan kata-kata bernada positif/negatif, emotikon, serta konteks kalimat. Untuk menjaga konsistensi, peneliti melakukan pemeriksaan ulang secara bertahap dan menerapkan kriteria pelabelan yang seragam selama proses berlangsung.Contoh hasil labelling sentimen positif, netral, dan negatif dapat dilihat pada Tabel 2.

Tabel 1. Contoh Hasil Labelling Data

Comment	***Label***
aku pakai yang glow cushion wardah bagus banget	2
bih baik makeover atau wardah iya bingung	1
aku tdinya pakai cream saya ganti jadi pakai wardah crystal secret ko jadi kusam padahal suka exfoliasi terus timbul bintik jerawat komedo kira kenapa iya	0

2.3 *Data Preparation*

Pada tahapan ini, terdapat lima proses utama yang dilakukan, yaitu *data cleansing*, *case folding*, *tokenizing*, *filtering*, dan *stemming*. Setelah seluruh data melalui kelima tahapan *preprocessing* tersebut dan berhasil dibersihkan dari *noise* seperti karakter khusus, tanda baca, serta kata-kata yang tidak relevan, data dianggap siap untuk digunakan dalam tahap selanjutnya. Langkah berikutnya adalah melakukan proses *vectorization* terhadap data yang telah dibersihkan, yang

kemudian dilanjutkan dengan pemodelan untuk analisis sentimen.

2.3.1. *Case Folding*

Case folding bertujuan untuk mengubah seluruh huruf kapital menjadi huruf kecil. Tahapan ini penting dilakukan untuk menyamakan representasi kata, sehingga kata yang memiliki makna sama tetapi ditulis dengan format huruf berbeda dapat dikenali sebagai satu entitas yang sama.

Tabel 2. *Case Folding*

Before Case Folding	***After Case Folding***
Aku pake yg glow cushion wardah, bagus bgt😍	**aku** pake yg glow cushion wardah, bagus bgt😍

Pada Tabel 3 dapat dilihat bahwa kata "Aku" yang sebelumnya ditulis dengan huruf kapital di awal, setelah dilakukan proses *case folding* menjadi "aku" dalam huruf kecil. Hal ini menunjukkan bagaimana case folding menyederhanakan format teks untuk kebutuhan analisis data selanjutnya.

2.3.2. *Cleansing Data*

Pada tahap data *cleansing*, dilakukan proses pembersihan teks komentar agar data yang digunakan dalam analisis berada dalam kondisi bersih dan siap untuk diproses lebih lanjut. Beberapa langkah pembersihan yang dilakukan meliputi penghapusan mention (karakter "@"), hashtag, tautan URL, angka, tanda baca, karakter non-alfabet, serta spasi berlebih. Proses ini dilakukan menggunakan fungsi *clean_text* yang diimplementasikan dengan bantuan modul *regex* dan *string* pada Python. Pada Tabel 4 terlihat bahwa tanda koma setelah kata wardah dan emotikon di akhir kalimat telah dihapus, sesuai dengan tujuan dari tahap pembersihan ini.

Tabel 3. *Cleansing Data*

Before Cleansing Data	***After Cleansing Data***
aku pake yg glow cushion **wardah,** bagus bgt😍	aku pake yg glow cushion **wardah** bagus bgt

2.3.3. *Filtering Data*

Tahapan *filtering* dilakukan untuk menghapus kata-kata yang tidak relevan dan tidak mendukung analisis sentimen. Selain itu, dilakukan penyaringan lanjutan terhadap komentar yang mengandung kata-kata tertentu yang dianggap tidak relevan, seperti "prelove", "jual", "edit", dan sebagainya. Kata-kata tersebut dikumpulkan dalam sebuah daftar yang disebut *banned_words*, dan digunakan dalam pola pencarian dengan regex untuk mencocokkannya terhadap isi kolom komentar. Dengan bantuan fungsi str.contains, baris komentar yang mengandung kata-kata tersebut kemudian dihapus dari dataframe. Hasil dari proses ini adalah dataset yang lebih bersih dan terfokus pada komentar-komentar yang relevan dengan topik analisis.

2.3.4. *Normalisasi Data*

Tahap normalisasi data bertujuan untuk menggantikan kata-kata yang memiliki variasi penulisan atau singkatan menjadi bentuk standar yang lebih umum. Untuk itu, digunakan kamus (*dictionary*) yang berisi pasangan kata slang dan bentuk standarnya. Dalam penelitian ini, proses normalisasi dilakukan dengan menggunakan file `*slang_words.txt*` berisi kata-kata slang, yang diadaptasi dari repositori GitHub NLP_bahasa_resources [10]dan kemudian dilengkapi dengan tambahan kata slang yang sering ditemukan dalam komentar pengguna.

Selanjutnya, fungsi *normalize_text* diterapkan ke setiap komentar dalam dataset. Pada Tabel 5 terlihat hasil dari proses normalisasi, di mana kata "pake" diubah menjadi "pakai", "yg" menjadi "yang", dan "bgt" menjadi "banget". Hasil akhir dari normalisasi ini adalah teks komentar yang lebih bersih, konsisten, dan siap untuk dianalisis.

Tabel 4. Normalisasi Data

Before Normalisasi Data	*After Normalisasi Data*
aku **pake yg** glow cushion wardah bagus **bgt**	aku **pakai yang** glow cushion wardah bagus **banget**

2.3.5. *Tokenizing Data*

Tokenizing bertujuan untuk memisahkan kalimat atau teks menjadi bagian-bagian kecil yang disebut token, yang dapat berupa kata, frasa, atau karakter. Proses ini penting untuk mempersiapkan data agar dapat dianalisis lebih lanjut pada tahapan *text mining* atau *machine learning*. Hasil dari proses *tokenizing* dapat dilihat pada Tabel 6.

Tabel 5. *Tokenizing Data*

Before Tokenizing Data	*After Tokenizing Data*
aku pakai yang glow cushion wardah bagus banget	aku,pakai,yang,glow,cushion,wardah,bagus,banget

2.3.6. *Stopword Removal*

Tahap *stopword removal* bertujuan untuk menghapus kata-kata yang tidak memiliki makna signifikan dalam analisis teks, seperti kata hubung, kata ganti, atau kata-kata umum lainnya yang sering muncul dalam teks tetapi tidak memberikan kontribusi penting terhadap analisis sentimen atau topik. Untuk proses ini, digunakan daftar kata-kata *stopword* yang diambil dari file `*stop_words.txt*` diadaptasi dari repositori GitHub NLP_bahasa_resources [10]. Daftar tersebut kemudian dilengkapi dengan tambahan *stopwords* yang sering muncul dalam komentar-komentar pengguna. Setelah daftar *stopword* terbentuk, fungsi`*remove_stopwords*` diterapkan untuk setiap teks dalam *dataset*. Fungsi ini akan memeriksa setiap kata dalam teks dan menghapusnya jika kata tersebut ada dalam daftar *stopwords*. Proses ini menghasilkan teks yang lebih bersih dan relevan, tanpa kata-kata yang tidak memiliki kontribusi penting. Hasil dari proses *stopword removal* dapat dilihat pada Tabel 7.

Tabel 6. Stopword Removal

Before Stopword Removal	*After Stopword Removal*
aku,pakai,**yang**,glow,cushion,wardah,bagus,banget	pakai,glow,cushion,wardah,bagus,banget

2.3.7. *Stemming Data*

Pada tahap stemming, proses ini dilakukan untuk mengubah setiap kata dalam komentar menjadi bentuk kata dasar, sehingga variasi kata yang memiliki makna serupa dapat diseragamkan. Dalam penelitian ini, proses stemming dilakukan menggunakan pustaka Sastrawi dengan mengimpor dan membuat objek stemmer dari kelas StemmerFactory. Selanjutnya, dibuat sebuah dictionary tambahan yang bernama corrections yang berisi pasangan kata yakni kata tidak baku atau turunan yang umum digunakan dalam komentar, dan bentuk dasarnya untuk memperbaiki hasil stemming yang belum sepenuhnya akurat oleh Sastrawi. Fungsi correct_stem() kemudian didefinisikan untuk menerapkan proses stemming sekaligus melakukan penggantian kata-kata sesuai dengan daftar corrections yang telah ditentukan. Hasil dari proses *stemming* dapat dilihat pada Tabel 8.

Tabel 7. *Stemming Data*

Before Stemming Data	*After Stemming Data*
nyobain,hijau,cocok,kulit,super,**berminyak**	nyobain,hijau,cocok,kulit,super, **minyak**

2.4. *Modelling*

Setelah tahap data *preparation* selesai, diperoleh sebanyak 878 data komentar yang telah dibersihkan dan siap digunakan dalam proses pemodelan. Data tersebut terdiri dari 425 komentar berlabel positif, 305 netral, dan 148 negatif. Tahap pemodelan diawali dengan proses pembobotan teks menggunakan metode TF-IDF untuk mengubah komentar ke dalam bentuk representasi numerik. Setelah proses transformasi, sebanyak 878 komentar digunakan dalam *modelling*. Data tersebut kemudian dibagi menjadi dua bagian, yaitu 80% atau sebanyak 702 data digunakan sebagai data berlabel, dan 20% atau sebanyak 176 data digunakan sebagai data tidak berlabel untuk proses *semi-supervised learning*. Data berlabel selanjutnya dibagi menggunakan teknik *5-fold cross-validation*, sehingga seluruh data secara bergantian digunakan sebagai data latih dan data uji. Untuk mengatasi ketidakseimbangan jumlah komentar pada masing-masing kelas sentimen, diterapkan teknik SMOTE pada data latih. Langkah ini dilakukan agar model tidak bias terhadap kelas yang dominan dan memiliki performa klasifikasi yang lebih seimbang.

Model awal dilatih menggunakan algoritma *Multinomial Naive Bayes*, setelah proses pelatihan awal, dilakukan pendekatan *semi-supervised learning* menggunakan teknik *pseudo-labeling*, di mana komentar tidak berlabel yang diprediksi dengan tingkat kepercayaan ≥ 80% secara otomatis ditambahkan kembali ke dalam data latih. Proses ini dilakukan secara iteratif untuk meningkatkan jumlah data pelatihan tanpa perlu melakukan pelabelan manual tambahan. Setelah tahap *pseudo-labeling* selesai, jumlah data pelatihan meningkat tergantung jumlah komentar tidak berlabel yang memenuhi ambang batas keyakinan. Model kemudian dilatih ulang dengan data gabungan tersebut, dan hasil evaluasi menunjukkan adanya peningkatan akurasi, *precision*, *recall*, dan *F1-score*, khususnya pada kelas positif dan netral. Contoh hasil prediksi sentimen dari komentar tidak berlabel beserta tingkat kepercayaannya disajikan pada Tabel 8.

Tabel 8. *Pseudo Labelling*

Komentar	Label	Confidence(%)
bagus vinyl maybeline glazed teman	0	0.8113
bagus npure please jawab butuh kulit jerawat sensitif	1	0.8746
micelar water favorit	2	0.8258

2.5. *Evaluation*

Evaluasi dilakukan untuk menilai performa model dalam mengklasifikasikan sentimen komentar sebelum dan sesudah diterapkannya metode *semi-supervised learning*. Pada tahap awal, model dilatih hanya dengan menggunakan data berlabel sebanyak 702 komentar, dan dievaluasi menggunakan teknik *5-fold cross-validation*. Hasil evaluasi awal ini menunjukkan rata-rata akurasi sebesar 0,7565, yang menjadi baseline dalam model ini. Setelah pelatihan awal selesai, metode *pseudo-labeling* mulai diterapkan. Komentar tidak berlabel sebanyak 176 data diprediksi oleh model awal, dan hanya komentar yang memiliki *confidence score* ≥ 80% yang dipilih untuk ditambahkan ke data latih. Proses ini dilakukan secara bertahap dalam empat iterasi. Pada setiap iterasi, model dilatih ulang dengan menambahkan data *pseudo-labeled* dari hasil prediksi sebelumnya. Evaluasi dilakukan kembali setelah setiap iterasi untuk mengamati sejauh mana penambahan data *pseudo-labeled* berpengaruh terhadap performa model.

Hasil evaluasi dari keempat iterasi ditampilkan pada Tabel 9. Rata-rata akurasi model meningkat secara bertahap dari 0,7565 di iterasi awal menjadi 0,7733 pada iterasi keempat, yang merupakan hasil terbaik. Iterasi dihentikan pada tahap keempat karena pada iterasi tersebut tidak ditemukan lagi komentar *pseudo-label* dengan tingkat keyakinan ≥ 80%, sehingga tidak ada tambahan data yang memenuhi syarat untuk dilatih ulang.

Tabel 9. *5 Fold Cross Validation*

Fold	Accuracy				
	Evaluasi Awal	Iterasi 1	Iterasi 2	Iterasi 3	Iterasi 4
1	0.7412	0.7841	0.7841	0.7955	0.7740
2	0.7811	0.7330	0.7159	0.7216	0.7514
3	0.7456	0.7386	0.75	0.7386	0.7955
4	0.7988	0.7543	0.7670	0.7841	0.7330
5	0.7160	0.7829	0.7943	0.7784	0.8125
Mean	**0.7565**	**0.7586**	**0.7623**	**0.7636**	**0.7733**

Selanjutnya, evaluasi per kategori sentimen dilakukan untuk mengukur metrik *Precision*, *Recall*, dan *F1-Score*. Untuk sentimen positif dapat dilihat pada Tabel 11, model mengalami penurunan *F1-Score* pada iterasi pertama yang disebabkan oleh turunnya nilai *Recall* dari 0.76 menjadi 0.75. Namun, setelah iterasi pertama, performa model kembali konsisten dengan nilai *F1-Score* stabil di angka 0.80. Pada iterasi keempat, model menunjukkan peningkatan performa dengan naiknya *Precision* menjadi 0.85 dan *Recall* menjadi 0.78, sehingga menghasilkan *F1-Score* tertinggi sebesar 0.81.

Tabel 10. Hasil Evaluasi Sentimen Positif

Iterasi	Sentimen Positif		
	Precision	***Recall***	***F1-Score***
Model Awal	0.84	0.76	0.80
1	0.84	0.75	0.79
2	0.84	0.76	0.80
3	0.85	0.76	0.80
4	0.85	0.78	0.81

Pada kategori sentimen netral yang ditampilkan pada Tabel 12, model menunjukkan konsistensi performa dengan *F1-Score* bertahan di angka 0.80 dari model awal hingga iterasi ketiga, kemudian meningkat menjadi 0.81 pada iterasi keempat. Pada iterasi pertama dan kedua, terjadi peningkatan nilai *Recall* dari 0.80 menjadi 0.82, sementara *Precision* tetap konstan di angka 0.79. Selanjutnya, pada iterasi keempat, *Precision* meningkat menjadi 0.81 sementara *Recall* tetap di angka 0.82. Kombinasi peningkatan ini menghasilkan *F1-Score* tertinggi sebesar 0.81.

Tabel 11. Hasil Evaluasi Sentimen Netral

Iterasi	Sentimen Netral		
	Precision	***Recall***	***F1-Score***
Model Awal	0.79	0.80	0.80
1	0.79	0.81	0.80
2	0.79	0.82	0.80
3	0.79	0.82	0.80
4	0.81	0.82	0.81

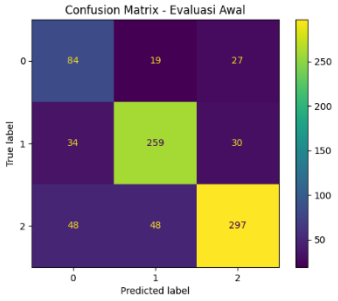
Sementara itu, untuk sentimen negatif dapat dilihat pada Tabel 12, performa model tergolong lebih rendah dibanding kategori lainnya. *F1-Score* mengalami sedikit peningkatan dari 0.57 pada model awal menjadi 0.59 pada iterasi keempat.

Tabel 12. Hasil Evaluasi Sentimen Netral

Iterasi	Sentimen Negatif		
	Precision	***Recall***	***F1-Score***
Model Awal	0.51	0.65	0.57
1	0.52	0.64	0.57
2	0.53	0.63	0.58
3	0.52	0.64	0.57

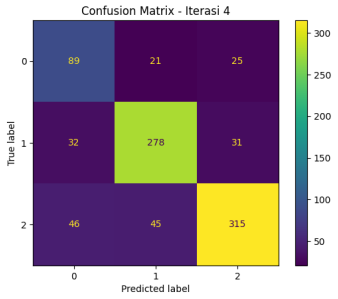
4	0.53	0.66	0.59

Lebih lanjut, analisis *confusion matrix* digunakan untuk mengevaluasi distribusi prediksi model terhadap label sebenarnya. Confusion matrix memberikan gambaran rinci mengenai jumlah prediksi benar dan salah pada masing-masing kelas. Pada evaluasi awal, menunjukan terdapat 84 komentar positif, 259 komentar netral dan 297 komentar positif yang diklasifikasikan secara benar. Hal ini terlihat pada Gambar 6 banyaknya komentar negatif yang justru diprediksi sebagai netral maupun positif.



Gambar 3. *Confusion Matrix* Evaluasi Awal

Setelah metode *semi-supervised learning* diterapkan, terjadi perbaikan bertahap, terutama dalam klasifikasi komentar negatif. Dapat dilihat pada Gambar 7 iterasi keempat menunjukan jumlah komentar positif yang diklasifikasikan secara benar meningkat dari 297 menjadi 315, sementara kesalahan klasifikasi ke kelas netral menurun dari 48 menjadi 45, dan ke kelas negatif dari 48 menjadi 46. Perbaikan juga terlihat dalam klasifikasi komentar netral, di mana prediksi benar meningkat dari 259 menjadi 278, meskipun tetap ada fluktuasi pada iterasi sebelumnya. Untuk sentimen negatif, meskipun peningkatannya tidak signifikan, prediksi benar tetap stabil dengan sedikit kenaikan dari 84 menjadi 89 pada iterasi keempat.

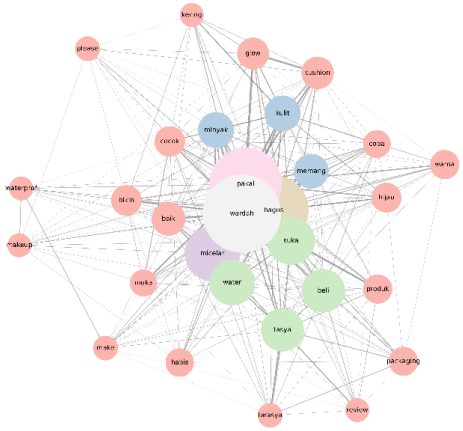


Gambar 4. *Confusion Matrix* Iterasi-4

2.6. *Deployment*

2.6.1. *Network Nodes* Sentiment Positif

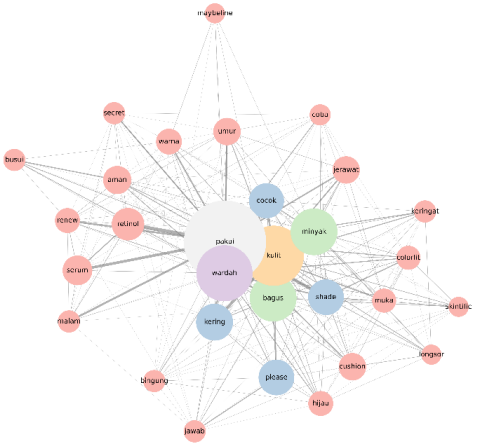
Gambar 8 menampilkan visualisasi network nodes yang merepresentasikan sentimen positif dari komentar pengguna TikTok terhadap unggahan Tasya Farasya dan Wardah. Kata seperti “pakai”, “bagus”, “wardah”, dan “suka” muncul sebagai node sentral berukuran besar, menandakan frekuensi tinggi dan keterkaitan kuat dengan kata lain. Hal ini menunjukkan bahwa produk Wardah, terutama micellar water, mendapat apresiasi positif. Kata-kata seperti “micelar”, “kulit”, “minyak”, dan “cocok” mengindikasikan bahwa pengguna merasa produk tersebut efektif dan sesuai dengan jenis kulit mereka. Sementara itu, kata “beli”, “produk”, “packaging”, dan “review” menunjukkan bahwa keputusan pembelian dipengaruhi oleh tampilan kemasan dan ulasan pengguna lain. Kemunculan kata “tasya” dan “farasya” mengisyaratkan adanya pengaruh dari beauty influencer terhadap persepsi konsumen. Secara keseluruhan, temuan ini menunjukkan bahwa sentimen positif dipengaruhi oleh pengalaman penggunaan, kecocokan produk, serta pengaruh eksternal seperti review dan rekomendasi publik figur.



Gambar 5. *Network Nodes Positif*

2.6.2. *Network Nodes* Sentiment Netral

Gambar 9 menampilkan visualisasi network nodes untuk sentimen netral dalam komentar pengguna TikTok terhadap unggahan Tasya Farasya dan Wardah. Kata seperti "pakai", "wardah", "kering", "kulit", dan "cocok" muncul dominan di posisi sentral. Meskipun bersifat netral, komentar cenderung berupa pertanyaan, keraguan, atau perbandingan, seperti ditunjukkan oleh kata "bingung", "please", dan "jawab". Kemunculan nama merek lain seperti "maybeline", "skintific", dan "colorfit" menunjukkan bahwa pengguna sering membandingkan produk Wardah dengan produk lain. Komentar pada kategori ini umumnya tidak menunjukkan kepuasan atau ketidakpuasan secara eksplisit, melainkan lebih fokus pada pencarian informasi atau pertimbangan sebelum membeli.

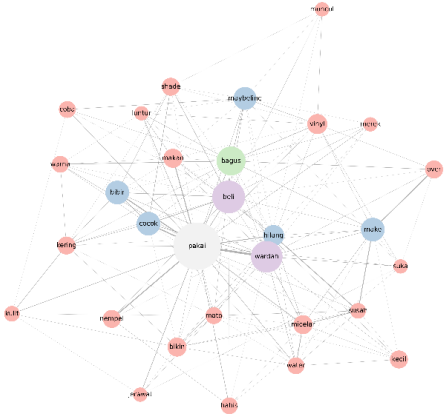


Gambar 6 *Network Nodes Neutral*

2.6.3. *Network Nodes* Sentiment Negatif

Gambar 10 menampilkan visualisasi network nodes yang merepresentasikan sentimen negatif dari komentar pengguna TikTok terhadap unggahan Tasya Farasya dan Wardah. Kata seperti "pakai", "hilang", "wardah", "beli", dan "bagus" muncul sebagai node sentral, meskipun "bagus" kemungkinan digunakan dalam konteks sarkastik atau perbandingan negatif. Node seperti "hilang", "luntur", "susah", dan "kecil" menunjukkan keluhan terkait performa produk, seperti mudah luntur atau kemasan terlalu kecil. Kemunculan kata "jerawat", "kering", dan "bikin" mengindikasikan efek negatif pada kulit. Sementara itu, nama merek lain seperti "maybeline" dan "vinyl" muncul dalam konteks perbandingan yang kurang menguntungkan bagi Wardah. Secara keseluruhan, komentar

negatif in mencerminkan ketidakpuasan terhadap kualitas, hasil, hingga harga produk.



Gambar 7 *Network Nodes Negative*

PEMBAHASAN

Penelitian ini menggunakan 2.716 komentar TikTok dari akun Tasya Farasya dan akun resmi Wardah dengan periode publikasi video antara tahun 2024 hingga 2025. Tujuannya adalah untuk menganalisis persepsi masyarakat terhadap produk Wardah berlabel "Tasya Farasya *Approved*", mengidentifikasi distribusi sentimen, serta mengevaluasi akurasi metode *Semi-Supervised Learning* (SSL) dalam klasifikasi sentimen. Setelah melalui proses pembersihan dan pelabelan data secara manual, diperoleh 878 komentar yang layak digunakan sebagai data analisis. Data ini dibagi menjadi dua bagian, yaitu 702 komentar sebagai data berlabel dan 176 sebagai data tidak berlabel. Seluruh data dikonversi ke bentuk numerik menggunakan metode Term Frequency- TF-IDF. Model dikembangkan menggunakan algoritma *Naive Bayes*, divalidasi menggunakan teknik *5-fold cross-validation*, dan diseimbangkan menggunakan SMOTE pada data latih untuk mengatasi ketimpangan distribusi kelas.

Hasil evaluasi menunjukkan bahwa akurasi model meningkat dari 0,7565 menjadi 0,7733 setelah empat iterasi. Nilai *F1-score* untuk sentimen positif dan netral meningkat hingga mencapai 0,81. Namun, peningkatan pada kelas negatif cenderung stagnan, hanya naik dari 0,57 menjadi 0,59. *Confusion matrix* menunjukkan bahwa model masih kesulitan dalam mengenali komentar bernada negatif secara akurat. Kondisi ini dipengaruhi oleh dominasi jumlah data pada kelas positif dan netral, yang secara signifikan lebih banyak dibandingkan kelas negatif. Meskipun SMOTE digunakan untuk menyeimbangkan distribusi saat pelatihan, teknik ini memiliki keterbatasan dalam konteks teks karena ia hanya mereplikasi data sintetik berbasis vektor numerik tanpa memahami konteks linguistik [11]. Akibatnya, kelas minoritas seperti sentimen negatif tetap tidak memiliki representasi yang cukup kuat dalam ruang fitur, sehingga model cenderung bias ke arah kelas mayoritas. Selain itu, karakteristik komentar TikTok yang cenderung pendek, informal, dan kontekstual turut menyulitkan model dalam mengidentifikasi ekspresi negatif secara tepat. Dengan demikian, meskipun pendekatan *semi-supervised learning* menunjukkan potensi dalam meningkatkan performa model di kondisi terbatas, hasil ini juga menegaskan perlunya peningkatan pada sisi pemodelan.

KESIMPULAN

Persepsi publik terhadap produk Wardah dengan label Tasya Farasya *Approved* menunjukkan kecenderungan yang positif, tercermin dari banyaknya komentar apresiatif, terutama terhadap produk *micellar water*. Pengaruh Tasya Farasya sebagai *influencer* turut membentuk persepsi tersebut, sejalan dengan temuan penelitian sebelumnya yang menunjukkan bahwa kehadiran Tasya Farasya dapat memberikan dampak positif terhadap keputusan pembelian konsumen, seperti pada produk Skintific di platform e-commerce TikTok [12]. Dari 878 komentar yang dianalisis, distribusi sentimen terdiri atas 425 sentimen positif, 305 netral, dan 148 negatif. Penerapan pendekatan *Semi-Supervised Learning* (SSL) dalam penelitian ini terbukti efektif dalam meningkatkan performa model, dengan akurasi yang meningkat dari 0,7565 menjadi 0,7733, serta peningkatan pada nilai *Precision*, *Recall*, dan *F1-Score*, khususnya untuk

kategori sentimen positif dan netral. Kontribusi utama dari penelitian ini terletak pada pemanfaatan SSL dalam analisis sentimen komentar TikTok, yang masih relatif jarang dieksplorasi, serta pada pengungkapan peran *influencer* dalam membentuk persepsi terhadap produk. Meski demikian, penelitian ini memiliki keterbatasan, di antaranya adalah dominasi data dari satu ketidakseimbangan data sentimen negatif yang masih sulit diatasi meskipun telah dilakukan penyeimbangan. Untuk pengembangan di masa depan, penelitian ini dapat diperluas dengan menambah variasi sumber data dari berbagai akun dan merek, menerapkan model SSL yang lebih kompleks, serta mengeksplorasi pendekatan lain.

REFERENCES

[1] I. E. Mufrida, "Orang Indonesia Paling Sering Habiskan Waktu untuk Main Sosial Media," GoodStats. Diakses: 22 Juni 2025. [Daring]. Tersedia pada: https://goodstats.id/article/orang-indonesia-paling-sering-habiskan-waktu-untuk-main-sosial-media-nETfh

[2] M. D. Daryanti, R. Hidayat, S. Rahayu, dan R. Fitrianty, "Pengaruh Strategi Pemasaran Digital Marketing, Word of Mouth dan Influencer Marketing terhadap Keputusan Pembelian melalui Minat Beli pada Produk Kecantikan MS Glow di Kota Sumenep," *Juremi: Jurnal Riset Ekonomi*, vol. 4, no. 6, hlm. 1321–1336, Mar 2025.

[3] D. Pertiwi, E. J. Hasibuan, dan A. Auza, "The Effect of Tasya Farasya's Youtube Account Vlog Content on Interest in Buying Subscribers," *Ilmu Komunikasi (JIPIKOM)*, vol. 4, no. 1, hlm. 34–43, Jun 2022, doi: 10.31289/jipikom.v4i1.1149.

[4] W. P. Sari *dkk*., "Pengaruh Endorsement 'Tasya Farasya Approved' Terhadap Brand Awareness Produk Somethinc Calm Down Cream (Survei Followers Instagram @tasyafarasya)," *Journal Innovation in Education (INOVED)*, vol. 2, no. 2, hlm. 309–318, Jun 2024.

[5] P. Oktaviani, F. S. P. W, N. M. A. Naafi'ah, dan S. Zuhri, "Peran Label 'Tasya Farasya Approved' Dalam Keputusan Pembelian Brand Skintific," *JUSKOP: Jurnal Studi Komunikasi dan Politik*, vol. 1, no. 2, hlm. 280–287, Des 2024.

[6] E. Kurniawati dan S. R. Rindrayani, "Pendekatan Kuantitatif dengan Penelitian Survei : Studi Kasus dan Implikasinya," *SOSIAL: Jurnal Ilmiah Pendidikan IPS*, vol. 3, no. 1, hlm. 65–69, Jan 2025, doi: 10.62383/sosial.v3i1.596.

[7] P. Arsi dan R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 1, hlm. 147, Feb 2021, doi: 10.25126/jtiik.0813944.

[8] E. Hasibuan dan E. A. Heriyanto, "Analisis Sentimen pada Ulasan Aplikasi Amazon Shopping di Google Play Store Menggunakan Naive Bayes Classifier," *Jurnal Teknik dan Science*, vol. 1, no. 3, hlm. 13–24, Okt 2022, Diakses: 23 Mei 2025. [Daring]. Tersedia pada: https://journal.admi.or.id/index.php/JTS/article/view/434/486

[9] A. Ligthart, C. Catal, dan B. Tekinerdogan, "Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification," *Appl Soft Comput*, vol. 101, Mar 2021, doi: 10.1016/j.asoc.2020.107023.

[10] L. Owen, "NLP bahasa resources," GitHub. Diakses: 23 Mei 2025. [Daring]. Tersedia pada: https://github.com/louisowen6/NLP_bahasa_resources/blob/master/combined_slang_words.txt

[11] M. Bystroński, M. Hołysz, G. Piotrowski, N. V. Chawla, dan T. Kajdanowicz, "SMOTExT: SMOTE meets Large Language Models," *Computation and Language*, vol. 1, Mar 2025, doi: https://doi.org/10.48550/arXiv.2505.13434.

[12] E. Rustiawati, I. Esther, L. N. Anargya, N. M. Kasyfan, dan S. S. Maesaroh, "Pengaruh Kredibilitas Dan Relevansi Beauty Influencer Terhadap Keputusan Pembelian Produk Skintific Di Tiktok @Tasyafarasya," *JAMBURA*, vol. 7, no. 2, 2024, [Daring]. Tersedia pada: http://ejurnal.ung.ac.id/index.php/JIMB